Science Library APR 23 1947

. 31, No. 1 February, 1947

Journal of Applied Psychology

DITED BY: DONALD G. PATERSON, UNIVERSITY OF MINNESOTA

Consulting Editors

AUL S. ACHILLES, *Psychological Corporation;* WALTER V. BINGHAM, *A.G.O., War Department;* AROLD E. BURTT, *Ohio State University;* ARTHUR I. GATES, *T. C. Columbia University;* OHN G. JENKINS, *University of Maryland;* IRVING LORGE, *T. C. Columbia University;* UINN MCNEMAR, *Stanford University;* WILLARD C. OLSON, *University of Michigan;* MES P. PORTER, *Swarthmore, Pennsylvania;* EDWARD K. STRONG, JR., *Stanford University;* ORRIS S. VITELES, *University of Pennsylvania;* JOSEPH ZUBIN, *N. Y. Psychiatric Institute.*

Table of Contents

dustrial Use of the Ishihara Tests for Color Blindness: C. E. JURGENSEN ... 1

New Method for Reading the Employment Questionnaire: W. F. WOOD ... 9

fferences in Interests Among Public Administrators: E. K STRONG, JR. 18

eliability and Validity of the Kopas Wage Earner Battery of Tests: B. BAXTER 39

uckiesh and Moss on Reading Illumination: A. C. HOFFMAN 44

me Evidence on the Validity of the Cardall Test of Practical Judgment: L. F. SMITH AND E. B HOGADONE 54

terests and General Educational Development: A. FRANDSEN 57

he Shipley-Hartford Retreat Scale as a Measure of Intellectual Impairment for Military Prisoners: M. P MANSON AND H. M. GRAYSON 67

he Special Leadership Training Program at the Signal Corps Officer Candidate School: R. E TAUBMAN 82

Note on Popular Pseudo-Psychological Beliefs in 1923 and in 1946: H. P. LONGSTAFF 91

pplied Research That Might Have Been: K. E. CLARK 94

ook Reviews 99

w Books, Monographs, and Pamphlets 107

rratum 109

ublished Bi-monthly by The American Psychological Association, Inc.
Prince and Lemon Sts., Lancaster, Pa., and
1515 Massachusetts Ave., NW, Washington 5, D. C.

ered as second-class matter, August 19, 1943, at the post office at Lancaster, Pa., under the Act of March 3, 1879

Copyright, 1947, by The American Psychological Association, Inc.

PUBLICATIONS OF

The American Psychological Association, Inc.

PSYCHOLOGICAL REVIEW

Editor: HERBERT S. LANGFELD, *Princeton University*

Contains original contributions of a theoretical nature; bi-monthly.
Subscription: $5.50 (Foreign $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

Editor: LYLE H. LANIER, *Vassar College*

Contains critical reviews of books and articles and critical and analytical summaries of psychological fields or subject matter; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL MONOGRAPHS

Editor: JOHN F. DASHIELL, *University of North Carolina*

Contains longer researches and laboratory studies which appear as units; published at irregular intervals at a cost to author of about $2.50 a page; author receives 150 copies gratis.
Subscription: $6.00 per volume of about 350 pages (Foreign $6.30). Single copies, price varies according to size.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

Editor: GORDON W. ALLPORT, *Harvard University*

Contains original contributions in the field of abnormal and social psychology, reviews and case reports; quarterly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editor: FRANCIS W. IRWIN, *University of Pennsylvania*

Contains original contributions of an experimental character; bi-monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

Editor: C. M. LOUTTIT, *Sampson College*

Contains noncritical abstracts of the world's literature in psychology and related subjects; monthly.
Subscription: $7.00 (Foreign $7.25). Single copies, $.75.

JOURNAL OF APPLIED PSYCHOLOGY

Editor: DONALD G. PATERSON, *University of Minnesota*

Contains material covering applications of psychology to business, industry, education, etc.; bi-monthly.
Subscription: $6.00 (Foreign $6.50). Single copies, $1.25.

APPLIED PSYCHOLOGY MONOGRAPHS

Editor: HERBERT S. CONRAD, *College Entrance Examination Board, Princeton*

Contains longer researches and studies in the field of applied psychology; published at irregular intervals at a cost to author of about $2.00 a page.
Single copies only, price varies according to size.

JOURNAL OF CONSULTING PSYCHOLOGY

Editor: LAURANCE F. SHAFFER, *Teachers College, Columbia University*

Contains articles in the field of clinical and consulting psychology, counseling and guidance; bi-monthly.
Subscription: $3.00 (Foreign $3.50). Single copies, $.60.

THE AMERICAN PSYCHOLOGIST

Editor: DAEL WOLFLE, *American Psychological Association*

Contains all official papers of the Association and articles concerning psychology as a profession; monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY

Editor: CALVIN P. STONE, *Stanford University*

Contains original contributions in the field of comparative and physiological psychology; bi-monthly.
Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

Subscriptions are payable in advance and are terminated at expiration.
Make checks payable to the American Psychological Association, Inc.

Subscriptions, orders, and other business communications should be sent to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 MASSACHUSETTS AVENUE N.W., WASHINGTON 5, D.C.

Journal of Applied Psychology

Vol. 31, No. 1 February, 1947

Industrial Use of the Ishihara Tests for Color Blindness

Clifford E. Jurgensen

Minneapolis Gas Light Company

Color blindness among American white males is estimated to be approximately six to eight per cent. Therefore, tests for color blindness are of interest not only to those industries in which normal color perception is a major part of production (such as textiles, paper, plastics, apparel, etc.) but to all industry. Every company employs office workers dealing with visible files utilizing colored tabs and guides, or maintenance men dealing with color-coded pipe or wire, or other employees concerned with color in small but not minor ways.

Merely because color discrimination appears to be involved in a particular job is no proof that color blindness is necessarily a handicap. Only direct measurement of successful employees on the job can determine whether normal color vision is really essential. In this respect tests for color blindness do not differ from tests in other fields. Miles and Craig[1] found, for example, that 7.2% of 375 male drygoods salesmen were color blind. Only one-half of these individuals were found to be a detriment to their respective firms, and the remainder included some who claimed to be "color experts." Miles and Craig suggest that "Perhaps every well-organized haberdashery should have its one or two high salaried color-blind employees to use generally throughout the store in matching brightness when the chromatic elements are so strong as to make this difficult for the ordinary person. Brightness matching is secondary to chroma matching, but it is just the additional part that when well done makes the combination perfect." [2]

Unpublished research of the author indicated this same situation in a paper mill manufacturing colored specialty papers. The most proficient color mixer was found to be color blind, although neither he nor his supervisor were aware of the fact. Specifications used by the color mixer

[1] Miles, W. R., and Craig, H. Color blindness in drygoods salesmen. *Person. J.*, 1931, 9, 437–449.

[2] *Op. cit.*, page 438.

invariably stated what color paper was being produced, and all containers of dye color were clearly labled regarding their contents. Consequently no chroma confusion could occur. The accuracy of the color match was therefore concluded to depend primarily on brightness. This single case does not prove, of course, that color blindness is a prerequisite for optimum color matching. It is possible that the particular individual involved was proficient in spite of rather than because of his color blindness. However, its relationship to the conclusion reached by Miles and Craig should not be overlooked, and it illustrates the principle that all tests should be validated regardless of their apparent relationship with job success.

The Ishihara Tests for Color Blindness have achieved considerable popularity in business and industrial situations. The test is most frequently administered by psychologically untrained testers. They must follow the Ishihara instructions which, though satisfactory for a psychologist, have been found by the author to be inadequate for lay testers.

The Ishihara instructions fail to emphasize the importance of standardized administration. In an industrial situation the test was used as a group test with approximately twenty employees in the group. The test administrator stood at one end of the room, the distance between the individuals and the test plates varied from three to fifteen feet, and many employees saw the plates only from an angle, In a factory situation the employees were tested at their work benches under artificial lights varying in type and intensity. In still another industrial situation the workers were allowed to look at the plates for as long a time as they desired. With such unlimited time, many saw more than one number and thus some color blind individuals gave the responses of normal individuals and vice versa.

The instructions are also inadequate for test interpretation by the lay person. In a sample of 100 industrial employees, the author found disagreement between psychologists and untrained testers in thirty per cent of the cases. Percentage interpretations were particularly common, employees having various types of color blindness being classified as having a given percentage of color normalcy. One administrator was found whose interpretations consistently disagreed with those of psychologists. Later it was found that this administrator lacked normal color perception himself, and interpreted responses of others according to agreement or disagreement with his own responses. The factory in which this test administrator was personnel superintendent had discarded the Ishihara test as being worthless!

In the field of psychological testing valid results can be obtained only if the tests are properly administered and interpreted. The factory super-

visor or personnel man, most often the test administrator in the industrial situation, can administer and interpret a test properly only when supplied with adequate instructions phrased in vocabulary and terminology understandable to him. In the scoring keys and revised instructions for the Ishihara test developed by the author, the terminology is admittedly scientifically loose. Terms such as "dark purple" and "light red" may seem objectionable to the psychologist, but it was found that psychologically naive individuals who used these instructions and keys obtained results which agreed with those secured by psychologists.

The scoring keys and revised instructions which follow pertain to the *Seventh Complete Edition of the Ishihara Tests for Color Blindness* inasmuch as this appears to be the one most frequently used in the industrial situation. However, few changes are required to adapt the instructions to other editions. For example, by omitting some of the items and renumbering the remaining plates the instructions and keys pertain to the abridged edition of the test.

Instructions for Using Ishihara Tests for Color Blindness

7th Edition
Complete edition of 32 plates

Procedure for Administering the Tests

1. The tests are intended to be used under normal daylight conditions. If the room is dark or if usual artificial lights are used, the results will be invalid. Inasmuch as adequate daylight conditions are not always obtainable when desired, and even so-called "normal" daylight fluctuates widely, a portable *daylight bulb* apparatus may be used which provides lighting conditions which remain uniform from one test situation to another while eliminating the errors introduced by other kinds of artificial lights.
2. Place the closed book containing the tests on the edge of a table.
3. Have the person being tested stand so that he looks directly down on the test plates.
4. Be sure that the distance between the plates and the eyes of the person being tested is as close as possible to 30 inches. This distance requirement will be fulfilled for most persons by placing the test on a table of standard height.
5. Be sure that the person being tested looks squarely at the plates; i.e., does not look at them from an angle.
6. Open the book so the person being tested sees plate number 1. Read the following to the employee:

   "On each page of this book there is a circular plate made up of various colored dots. One or two numbers can be seen in most of the plates, although in some plates you will see nothing other than unrelated dots. Numbers may be all one color or may be made up of several colors. You will be allowed three seconds to look at each plate. You are to tell me what you see at first glance."
7. Use plates 1 and 32 as examples of what the person being tested must do. Inasmuch as these two plates can be read by all persons, it does not matter if the person tested looks at these two plates for more than three seconds.

8. Do not give leads by questions or otherwise. (For example, do not ask, "Do you see another number?")
9. Do not tell the person being tested whether he is giving the right or wrong answers until the test has been completed.
10. Show plate 32 (the last page) before plate 26. The reason is that all persons can read number 32, and thus will clearly understand the instruction, "Trace the line between each X." Failure to show plate 32 before plate 26 may result in confusion to the person being tested and inaccurate results from plates numbered from 26 to 31, inclusive.
11. Be sure that tracing on plates 26 to 32 is done with some object such as a brush, which will not mark the plates in any way.

Precautions

1. Plates must be kept out of the sunlight when not in use in order to prevent a gradual change in coloring due to exposure to light.
2. The order of showing the plates should be changed when, and only when, there is evidence that the person being tested is acquainted with the test or has memorized correct responses.
3. Care should be taken that the plates do not become pencil or finger marked, scratched, or defaced in any other way.

Use of the Scoring Key

By means of the scoring key a written record may be obtained of the kind and degree of color blindness present in any individual to whom the test is given. The scoring key (Table 1) consists of eight columns as follows:

Column 1 lists the plate numbers in the order they appear in the book. Each plate can be identified by the small numbers printed on the margin of the plate.
Column 2 lists all probable responses which will be made to each plate.
Columns 3–8 indicate the interpretation of each probable response. For example, if on plate 2 a person sees an 8, he has normal vision; if he sees a 3, he is red-green color blind; and if he sees no number, he is totally color blind.

The scoring key is used by having the person giving the test put a circle around each X which is opposite the number reported seen. For example, if the person being tested sees an 8 in plate number 2, a circle should be placed around the X in the "normal" column. If a 3 is seen, the X in the "red-green blind" column should be encircled. If no number is seen, a circle should be placed around the X in the "total color blind" column.

A note should be made of any cases in which the numbers are seen with great difficulty.

If a person sees a number which is not listed on the scoring sheet, that number should be recorded on the scoring key.

A record should be made of any mixed numbers. For example, it is possible, though improbable, that a person will see a 37 on plate number 5 on which normal persons see 57 and red-green color blind persons see 35.

Interpretation of Results

Common sense must be used when interpreting results. For example, if a person can see no number on plate number 20, it may mean that he is either normal or totally color blind. Obviously, he is not both. Which of the two interpretations should be considered as correct depends on responses made to other plates.

Table 1

Scoring Key for Ishihara Color Tests
(7th Edition)

Person Tested.....Tested by.....Date.....

Plate Number	Probable responses	Normal	Total color blind	Red-green blind	Red blind	Green blind	Incomplete red-green blind
2	8	X		X			
	3		X				
3	6	X		X			
	5		X				
4	29	X		X			
	70		X				
5	57	X		X			
	35		X				
6	5	X		X			
	2		X				
7	3	X		X			
	5		X				
8	15	X		X			
	17		X				
9	74	X		X			
	21		X				
10	2	X	X	X			
11	6	X	X	X			
12	97	X	X	X			
13	45	X	X	X			
14	5	X	X	X			
15	7	X	X	X			
16	16	X	X	X			
17	73	X	X	X			
18	5	X	X	X			
19	2	X	X	X			
20	45	X	X	X			
21	73	X	X	X			
22	26	X					X
	6				X		
	2					X	
23	42	X					X
	2				X		
	4					X	
24	35	X					X
	5				X		
	3					X	
25	96	X					X
	6				X		
	9					X	
26	Line	X		X	X		
27	Line	X		X	X		
28	Line	X		X	X		
29	Line	X		X	X		
30	Line	X	X	X			
31	Line	X	X	X			

Additional Comments:

Plates number 1 and 32 can be read correctly by all persons, and are used to make sure that the person being tested knows what he is supposed to do.

It is relatively simple to interpret the results if a person's color blindness or normality is clear cut. However, the tests can also be used to detect color weakness. For example, plates 2, 3, 4, and 5 are of the same type. If a person gives the normal response for three of these plates and the red-green blind response for the fourth plate, he is shown to have a red-green weakness but he should not be classed as red-green color blind.

The following tests may be considered as identical in type, and on the scoring key the various types are separated by asterisks (*****):

2–3–4–5
6–7–8–9
10–11–12–13
14–15–16–17
18–19–20–21
22–23–24–25
26–27
28–29
30–31

Color weakness is also indicated if the response consists of a mixed number, that is, if a double number is reported with one digit as seen by a person with normal color vision and the other digit as seen by a color blind individual.

As is important in the use of other tests, each company should establish its own standards of performance on the Ishihara Tests. Some companies may find that incorrect responses to one or two plates is serious and other companies may find the opposite. In setting up such standards, consideration should be given to the type of plate read inaccurately, the type of color blindness indicated, importance of mixed numbers, etc. It is inadequate to consider merely the number or per cent of plates missed.

Some persons will be able to see the number given by an individual with normal color perception and also that given by one who is color blind. This may or may not indicate color weakness. Many test administrators with considerable Ishihara experience are able to "see" numbers read by those who are color blind as well as those normally read. Occasionally a person will be found who though unfamiliar with the test will analyze the dots and so see more than one number. This is largely overcome by limiting the time during which he is permitted to look at a test plate, but may nevertheless occur. In such case a rare individual will be found whose color perception is normal but who will respond with numbers seen by color blind individuals with the belief that the other numbers are so obvious that the test must be a "trick."

The supplementary scoring sheet (Table 2) is especially intended for persons mentioned above. Four specified pairs of plates are compared on the basis of ease of reading. Specified hues are then compared and the individual tested is asked if he can see any difference between them. This is done by asking him to name the colors. Correct naming of colors automatically assures differentiation. If colors are named incorrectly, the individual is asked what difference, if any, he sees. Finally, the individual is asked which of two specified hues is the more vivid or is the "brightest." Hues listed on the supplementary scoring sheet refer to those used in the Ishihara tests and can be found in several of the test plates. The basic assumption follows the usual physiological theory that red-green blind individuals see those colors as gray, and so report them as less vivid in hue than blues and yellows which are seen as such.

The supplementary scoring sheet will seldom be needed, and is intended for only those cases where usual test administration leaves doubt as to the kind or

Table 2

Supplementary Scoring Sheet for Ishihara Color Tests

(7th complete edition)

Note: For use only when the regular scoring key leaves doubt as to the kind or degree of color blindness.

Person Tested.....Tested by.....Date.....

****

In each of the following pairs, which plate is easier to read?
(Check answer given.)

Number	Interpretation
10 or 19	10 normal—19 color weak
13 or 20	13 normal—20 color weak
14 or 18	14 normal—18 color weak
17 or 21	17 normal—21 color weak

****

Determine if the person being tested can differentiate between two dots and name the color of each dot. Make the following comparisons and record the responses. Under "Differentiate," write *yes* or *no*. Under "Name first color" and "Name second color," make a check mark if color is correctly named, otherwise write in the color given.

Comparisons	Differentiate	Name 1st Color	Name 2nd Color
Light green and dark green (plate 2)	.....	.....	.....
Light purple and dark purple (plate 2)	.....	.....	.....
Light orange and dark orange (plate 2)	.....	.....	.....
Dark green and dark purple (plate 2)	.....	.....	.....
Dark orange and dark green (plate 2)	.....	.....	.....
Dark orange and dark purple (plate 2)	.....	.....	.....
Dark red and dark purple (plate 22)	.....	.....	.....
Light red and light purple (plate 22)	.....	.....	.....
Light gray and light red (plate 22)	.....	.....	.....
Light gray and light purple (plate 22)	.....	.....	.....
Black and dark purple (plate 22)	.....	.....	.....
Black and dark red (plate 22)	.....	.....	.....

****

The following colors should be compared. Place a check mark after the color in each pair which appears to be the more vivid:

Bright red.....Bright blue.....
Bright red.....Bright yellow.....
Bright green.....Bright blue.....
Bright green.....Bright yellow.....
Light red.....Light blue.....
Light red.....Light yellow.....
Light green.....Light blue.....
Light green.....Light yellow.....

Interpretation: Individuals who are red-green color weak will say that blue and yellow are much more vivid than green and red.

degree of color blindness. Its use will usually be limited to cases where mixed numbers are reported (for example, 20 in plate 4) incorrect numbers are reported (for example, 2 instead of 7 as in plate 15), or when an individual reports seeing both the number seen by normal persons and that seen by color blind persons (for example 8 and 3 in plate 2).

Publication of these instructions and scoring keys for lay testers is not to be construed as indicating that the author advocates test administration by lay persons. Although this procedure may be relatively satisfactory in the case of some tests (provided adequate instructions are available and are followed closely), other tests can be properly administered and interpreted only by persons having a thorough background in psychometrics. Nevertheless, the fact remains that many business and industrial concerns do engage in testing without having adequately trained or supervised testers. Numerous controls are possible, and are used with varying degrees of success by different companies. Instructions which can be used by lay testers do not give a panacea, but they will permit better test results when used by psychologically naive individuals and so help pave the way for additional and more adequate test controls such as the employment of an adequately trained psychologist to supervise the test program.

Received February 25, 1946.

A New Method for Reading the Employment Questionnaire *

Wendell F. Wood

Division of Applied Psychology, Purdue University

Two areas of research in the personnel field have developed to the extent that today every employment man could have available information and methods that can take much of the guesswork out of his difficult job of employee selection and placement.

The first accomplishment has been in job analysis. *The Dictionary of Occupational Titles* has succeeded in defining 17,452 separate jobs which have brought for the first time uniformity of approach to the classification of occupations and full coverage of the general employment field in the United States. Following this many businesses and industries have proceeded to study their own job requirements and have arrived at complete job descriptions for their entire organizations. For an industry to have done this indicates that considerable progress has been made in modernizing the employment office and in accepting the basic principles of job determination by more scientific methods.

The second accomplishment has been made possible by the many applied psychologists who have sought more accurate methods of measuring the abilities of man. Each year finds more industries adding standardized tests to their employment procedure. Carefully chosen tests are invaluable aids in employee selection and placement.

It appears to the writer that a far greater use can and should be made of the available developments in the employment field. This paper suggests a method by which this can be done with a minimum of expense.

Based on the premise that a more efficient and effective employment job can be done by objectively analyzing the job requirements and comparing them with the applicant's qualifications, the following method is suggested for doing this by the simple mechanical means of using a prepared template for each job. This method has the advantage of great flexibility; it can be as simple or as comprehensive as one wishes to make it, and can be used with or without occupational or other test results.

* This article is a "prior publication," the author paying complete costs. The scheduled 80 pages per issue is thereby increased by the corresponding amount; thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted and printed in their regular turn.

A suggested method that can be used in setting up this employment and selection program is outlined in the following steps:

1. The first step is to prepare a listing of all occupations in the organization. The problem of classification can be easily solved by using that of the *Dictionary of Occupational Titles*.[1] It is also recommended that the same code number be used, for as time goes on the *Dictionary* is being accepted and used more widely in many industries. The complete classification system is given below although Group 3 is not applicable to the majority of industries.

OCCUPATIONAL CODES [2]

0	Professional and marginal occupations	
	0–0 through 0–3	Professional occupations
	0–4 through 0–6	Semi-professional occupations
	0–7 through 0–9	Managerial occupations
1	Clerical and sales occupations	
	1–0 through 1–4	Clerical and kindred occupations
	1–5 through 1–9	Sales and kindred occupations
2	Service occupations	
	2–0	Domestic service occupations
	2–2 through 2–5	Personal service occupations
	2–6	Protective service occupations
	2–8 through 2–9	Building service workers and porters
3	Agricultural, fishery, forestry, and kindred occupations	
	3–0 through 3–4	Agricultural, horticultural, and kindred occupations
	3–8	Fishery occupations
	3–9	Forestry (except logging) and hunting and trapping occupations
4, 5	Skilled occupations	
6, 7	Semi-skilled occupations	
8, 9	Unskilled occupations	

2. The second step is to prepare complete personnel job descriptions for each of the various occupations under consideration. From a study of a given job description and from the results of a conference with the foreman of that job, and perhaps a discussion of it with the time-study and safety man, a list of employee specifications or requirements for the job is drawn up. This list, in its final form, must have the approval of the foreman or supervisor of the given job, and of the safety man. It is

[1] An employment questionnaire should only contain the occupational titles in its check list for jobs actually existing in a given company. Examples used in this paper probably will not be complete for any particular company or industry, or may be too lengthy for many firms.

[2] *Dictionary of Occupational Titles, Part II Titles and Codes*, United States Department of Labor, for sale by Superintendent of Documents, United States Government Printing Office, Washington, D. C., 1939, Pg. IX.

important to have the foreman's opinion that he would probably accept any new-hire for the job who met the minimum requirements as worked out in the above manner.

3. The third step is to devise an employment questionnaire for the applicant or employee. This must contain all of the information items common to all personnel requirements for the various jobs such as age, height, weight, training, experience, etc. For this purpose it is best to tally first all the different requirements from the job requirement lists set up in step two. The information for the questionnaire items is secured from several sources such as the employment application form, the interview, and the occupational test results if tests are used. The application form used for demonstration in this article is not given as an example to be used in any particular industry or company. An application form of the type developed by Kurtz [3] for the Life Insurance Sales Research Bureau might well be used.

4. The fourth and last step is to construct a template (which can be made from heavy paper or light cardboard) for each job. The template should be in the form of a folder approximately the same size as the employment application form with cut-outs which correspond with the requirements of a subject job. The template then shields out irrelevant material when placed over the questionnaire. Beside each cut-out is written the minimum requirement or acceptable limits of the job-requirement item appearing in that position. If a new employment application form is made, it is suggested that it be printed on good stock using both sides of the paper. The size should be standard letter size ($8\frac{1}{2} \times 11$ inch) which may be easily mailed and easily filed. One side should contain all personal information and the other side all occupational information. The template is then cut so that the front side of the folder corresponds with the front side of the application form giving personal information and the back side of the folder corresponds with the back side of the application form giving occupational information. This may be demonstrated by the following illustration: Figure 1 shows the front side of an application form, with all of the customary items of information required from the applicant. These are listed under the headings of Personal, Education, and Work History.

Figure 2 shows the employment questionnaire being inserted into the template folder. The job of operator lathe turret has been selected to describe how the template works.

Figure 3 shows the front side of the template folder superimposed on the corresponding side of the questionnaire. The items essential to the

[3] Kurtz, A. F. Selecting salesmen by personal history items. *Psychol. Bull.*, 1939, 36, 528.

ABC COMPANY
APPLICATION FOR EMPLOYMENT
(Please answer all questions by placing answers in spaces provided)

PERSONAL Date Oct. 1, 1946

Full Name John Paul Doe Social Security No. 62-6262-62

Present Address 100 North Street City Chicago State Ill.

Married Yes X No ___ Children 4 Dependents 5 Age 37

Sex Male ✓ Female ___ Height 6'1" Weight 205 Color White Nationality U.S.A.

Place of Birth Chicago, Ill. Date of Birth June 3, 1908

State of Health Excellent Disabilities none Glasses no

Are you willing to be examined by the company doctor? Yes

NAME Last Doe First John Middle Paul SOC. SEC. NO. 62-6262-62 DATE Oct. 1, 1946

EDUCATION

	Years	Name of school	Course or degree
Grammar School	8	Emerson	?
High School	4	Washington	graduated
College	0		
University	0		
Trade School	2	Chicago Tech	Machine Shop
Other	0		

WORK HISTORY

Please indicate the last two places where you have been employed:

Present or last employer X Y Z Company

His address 200 South Street His product Trucks

Date Employed: From 1936 to 1940 Total 4 Years

Your occupation or job there Engine Lathe Hourly or weekly wage 75¢

Next previous employer Square Circle Co

His address 400 West Street His product Machine Tools

Date employed: From 1940 to 1946 Total 6 Years

Your occupation or job there Turret Lathe Hourly or weekly wage $1.20

Signature of Applicant John Paul Doe Date Oct. 1, 1946

FIG. 1. Front side of an application blank.

job of a turret lathe operator, as listed in the job description, are: 2 years high school minimum, 2–3 years of trade training in machine shop work, plus 3 years experience on the turret lathe and 2 years experience on the engine lathe. Physical requirements for this job demand a large-type man, 30 to 40 years old. He should be at least 6 feet tall and weigh

190 to 240 pounds. We may say, in addition, and for the sake of example, that occupational research in the subject factory has determined that this type of machine operation requires a mental ability score of 100 to 110, a mechanical ability score above the 85th percentile, and a "personality" test rating of A or B. A married man is desired who has two or more children. Visual performance requirements for a vertical turret lathe operator were found to be for near vision: acuity, both eyes,

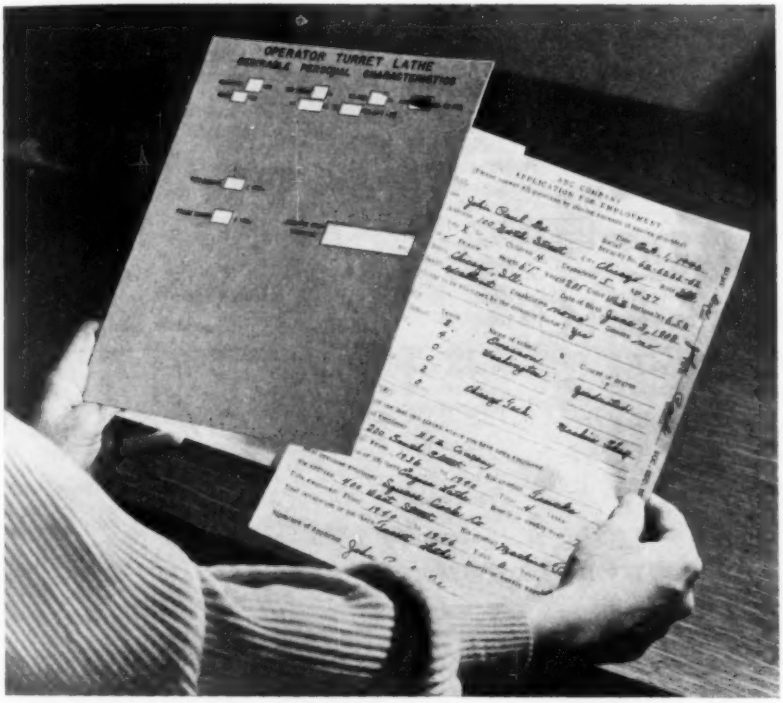
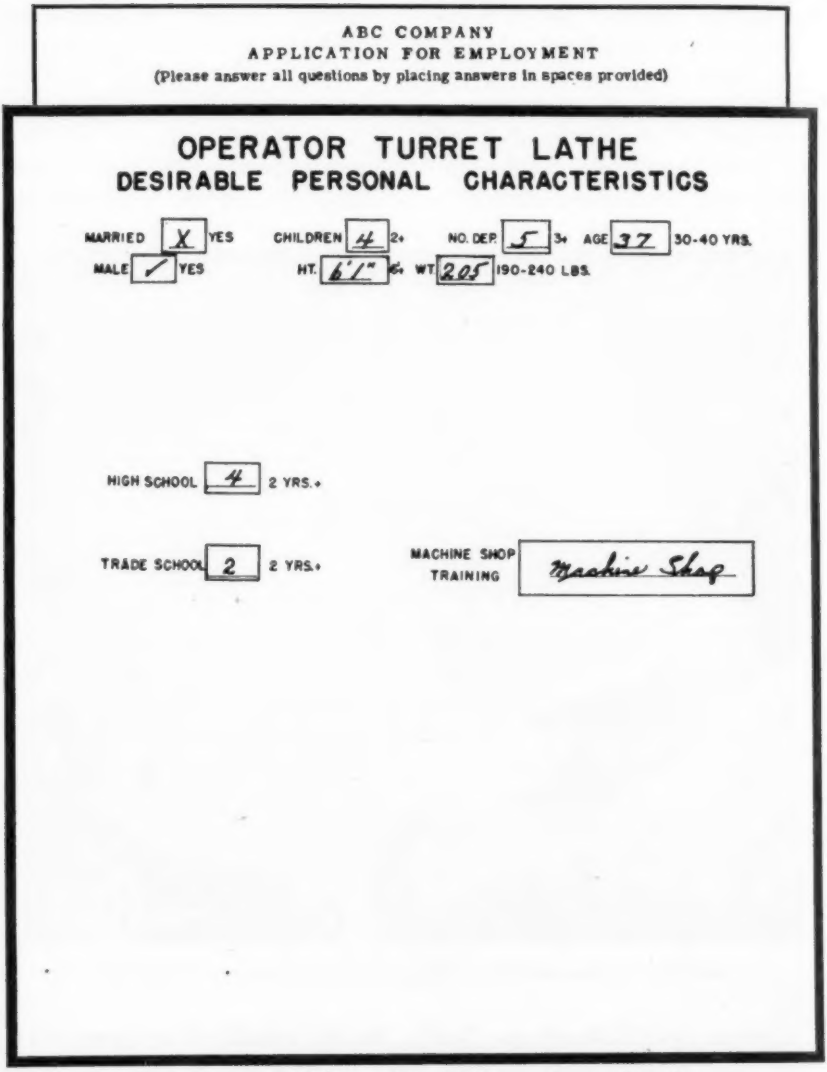


FIG. 2. Application blank being inserted into a template folder.

8–15; right eye, 7–15; left eye, 7–15. Phoria, vertical 1–7; lateral 6–11. (Far vision requirements, as with other possible test scores, have been left out for the sake of brevity.)

Figure 4 shows the back side of the application form on which the applicant is asked to list the position he is applying for, the years of experience he has had on that job plus all other specific job experience he has had. As may be seen, this is a comprehensive list in the four categories of shop or factory, trades, office-clerical, and professional. Blanks are left for possible write-in information. The box in the lower right hand


ABC COMPANY
APPLICATION FOR EMPLOYMENT
(Please answer all questions by placing answers in spaces provided)

OPERATOR TURRET LATHE
DESIRABLE PERSONAL CHARACTERISTICS

MARRIED [X] YES CHILDREN [4] 2+ NO. DEP. [5] 3+ AGE [37] 30-40 YRS.
MALE [✓] YES HT. [6'1"] 6+ WT [205] 190-240 LBS.

HIGH SCHOOL [4] 2 YRS.+

TRADE SCHOOL [2] 2 YRS.+ MACHINE SHOP TRAINING [Machine Shop]

FIG. 3. Front side of template folder superimposed on the corresponding side of an application blank.

corner contains space for the posting of occupational and visual test scores.

Figure 5 shows the back side of the template folder, which is also concerned with occupational information, superimposed on the corresponding, the back side, of the application form. Here the requirements,

OCCUPATIONAL INFORMATION

Position Applied For: Turret or Engine Lathe — Years of Experience: 10

Please check below the job or occupation you are best fitted for by reason of training or experience by indicating the years of experience you have had in each.

SHOP OR FACTORY		TRADES		OFFICE-CLERICAL		PROFESSIONAL	
Automatic Screw	1	Blacksmith		Bookkeeper		Accountant	
Boring Mill	1	Bricklayer		Cashier		Auditor	
Drill power		Carpenter		Clerk		Chemist	
Drill Radial		Electrician		General		Engineer	
Gear Cutter	2	Millwright		File		Chemical	
Grinder		Painter		Ledger		Civil	
Internal		Pipefitter		Mail		Electrical	
External	1	Plumber		Payroll		Industrial	
Hand Screw		Powerhouse		Sales		Metallurg.	
Lathe Engine	4	Heat				Mechanical	
Lathe Turret	6	Light		Messenger		Production	
Milling		Power		Office Machines		Sales	
Machine Repair		Refrigeration		Addressograph		Stationary	
Polisher		Air Condition		Calculator			
Tool & Die Maker		Welding		Comptometer			
Tool Inspector		Gas		Dictaphone		Public Relat.	
Tool Clerk		Electric		Duplicator		Ind. Relations	
Tool Grinder				I. B. M.		Personnel	
Tool Designer		Cook		Key Punch		Ind. Psychol.	
Tool Trouble		Baker		Tabulator		Social Work	
		Waiter				Physician	
Foundry Labor		Kitchen Helper		Receptionist		Surgeon	
Coremaker		Meat Cutter		Telephone Oper.		Reg. Nurse	
Moulder				Timekeeper		Training	
Melter		Guard		Stenographer			
Pour-Off		Watchman		Secretary		Draftsman	
Heat Treat		Fireman		Statistician		Architect	
		Elevator Oper.					
Pattern Maker				Collector			
Wood		Gardner		Interviewer			
Metal		Dock Hand		Investigator			
		Freight Handler		Salesman			
Clerk		Laborer					
Packing							
Receiving		Chauffeur					
Shipping		Truck Driver					
Stock	1	Trucker Elect.					
Traffic		Mechanic Auto					
Parts Inspector		Laboratory Ass.					
Quality Control		Photographer					
Time Study		Gang Boss					
Safety		Foreman					
		Supervisor					
		Manager					

PLEASE DO NOT WRITE IN THIS SPACE

Ment.	107	FAR		
Mech.	90% ile	Phoria	Vert.	5
Cler.			Lat.	9
Typ.		Acuity	Both	10
Int.			Rt.	9
Pers.			Lt.	9
		Depth - - -		9
		Color - - -		9
		NEAR		
		Acuity	Both	8
			Rt.	6
			Lt.	6
		Phoria	Vert.	6
			Lat.	10

FIG. 4. Back side of an application blank.

as indicated by the job description for a turret lathe operator, may be checked. There were, it may be recalled, in addition to certain occupational and visual test requirements, 3 years experience on the turret lathe, and 2 years experience on the engine lathe. This back side of the template also has another valuable use. Frequently the employment

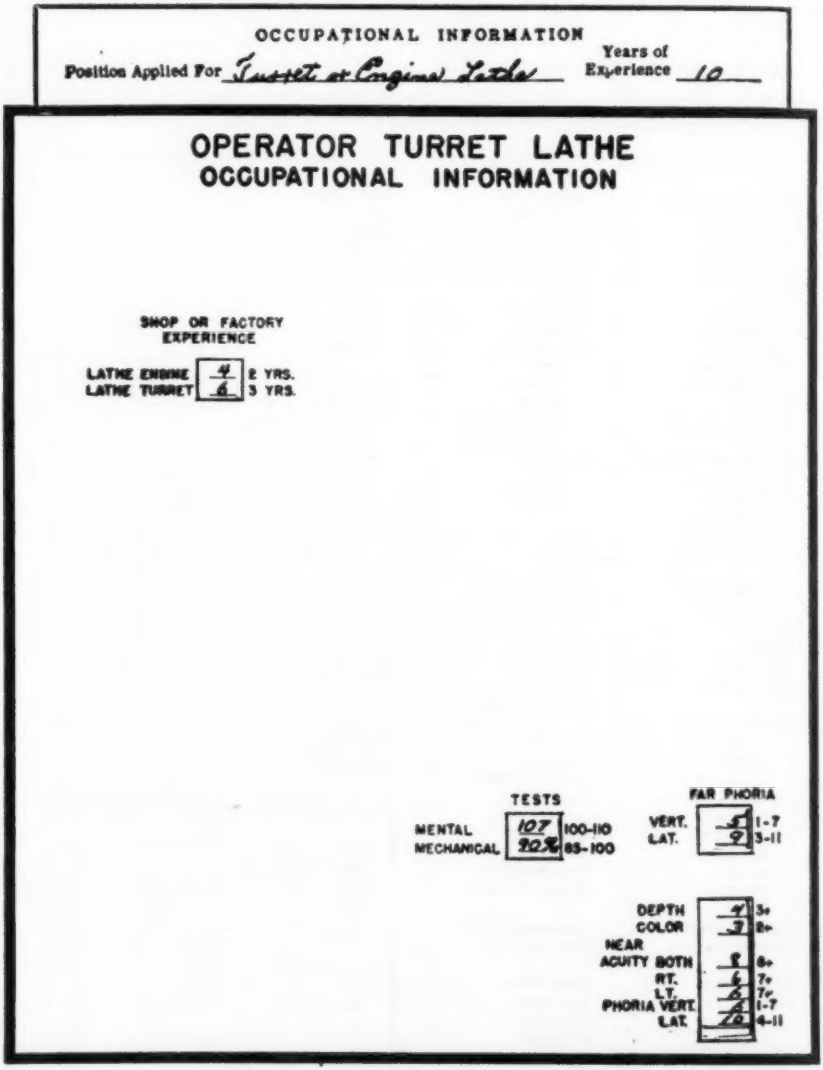


FIG. 5. Back side of template folder superimposed on the corresponding side of an application blank.

division has to find a person with a particular work background. For example, it may have a requisition for a stenographer who has had payroll, cashier, and I.B.M. experience. A special template can readily be prepared with cut-outs for these particular items. Any clerk can then be assigned the job of hunting through the personnel files where, with the

aid of the template, he will "pull" the likely application forms which seem to fill the requirements. The employment man can then take over and make a final selection after interviewing the prospects.

When the employment interviewer has finished the interview and has completed the applicant's employment questionnaire he finds it a very simple process to lay the proper template over it and study the essential minimum job requirements in terms of what the applicant has to offer.

It appears that this system of employee selection and placement has the following advantages:

1. There can arise no question between the employment office and the factory concerning job requirements, once they are set up with the cooperation of both.
2. An applicant can be judged solely by what he has to offer in terms of his qualifications for a particular job.
3. The employment man has definite specifications of the requirements he should look for on a given job.
4. The tendency to rationalize an applicant's qualifications for one job because of his personality traits or his outstanding experience not related to the subject job, is minimized.
5. Considerable time may be saved by the interviewer when he is obliged to hunt through a number of applications to find one with the requirements for a particular job.
6. The employment department, and particularly the interviewers, will learn to think more readily in terms of job elements and may find themselves generalizing and reading information into the application form as easily as in the past.

Received December 10, 1946.

Differences in Interests Among Public Administrators

Edward K. Strong, Jr.

Stanford University

Does the term "public administrator" refer to a more or less homogeneous group of men, or is it used to refer to the upper level jobs in the public service without implication as to their similarity?

In planning curricula for the training of young men and women for future public service it is important to know the answer to this question, to know whether there are actually different kinds of people in public administration.

In formulating policies as to the selection, training and promotion of people in civil service positions it is equally important to know whether they all belong to one general group or to several sub-groups.

Certain aspects of the subject have been considered in three previous articles. In the first article (**2**) it was shown that certain sub-groups of public administrators averaged relatively low on the public administrator interest scale, also certain individuals scored distinctly low on the scale. Evidence to be presented below shows that these sub-groups differ very considerably in their interests from the majority of public administrators. Evidence in the first article suggested that the individuals who scored so low are more professional men at heart than administrators, that is, they prefer to do the work themselves or with the assistance of a small staff of technical assistants rather than to direct a larger force engaged in a variety of activities.

In the second article (**4**) it was shown that the interests of senior public administrators differ appreciably from junior administrators, suggesting that men should be selected on a different basis for top positions than for the lower level administrative positions. A third article, confined to the interests of Forest Service Men, (**3**) substantiates the above.

In this article it will be shown that certain sub-groups of public administrators are very similar and that certain sub-groups differ greatly from one another in their interests.

The Data

The Committee on Public Administration of the Social Science Research Council obtained for the writer 550 Vocational Interest Blanks filled out by public administrators who were, in the judgment of the Committee, successful administrators. Some of these blanks for one

reason or another could not be used. Additional blanks, particularly of forest service men and city school superintendents, were supplied from our files.

The public administrators have been classified as far as possible according to the functions they perform, such as welfare, personnel, taxation, etc. Some of the administrators could not readily be classified in this manner for the reason that they direct employees engaged in many different activities, as for example, hospital superintendent, city manager, and senior administrator in the Department of Agriculture. Twenty-two sub-groups resulted but six of these were too small for statistical purposes and are not included here. The remaining sixteen sub-groups are listed in Table 1 with the number of cases in each.

Two questions concern us here: First, How do public administrators differ in their interests? and second, How much do they differ?

Classification of Public Administrators

A public administrator interest scale has been constructed in the same manner as other occupational interest scales based upon the interest records of 571 public administrators (**2**). On that scale 68.1 per cent of public administrators rate A, 15.8 per cent rate B+ and 9.4 per cent rate B. It is therefore possible to select a particular combination of specific activities in which most public administrators are interested. Certain sub-groups of public administrators average, however, appreciably below 50, the mean of all. Such sub-groups are office manager, 40; public health, 43.6; chemist-physicist, 44.4; city manager, 45.2; and engineer, 46.8.

Public administrators as a whole score high on the OL scale, which means they have the interests of the better paid business and professional men rather than the interests of common laborers. Their mean score on this scale is 60.7, sigma of 7.17. Comparable data for business and professional men are: lawyer, 64.4; president of a manufacturing concern, 63.4; sales manager, 63.3; personnel manager, 61.4; psychologist, 60.9; production manager, 60.2; accountant, 59.5; banker, 58.1; office worker, 57.0; musician, 53.8; caprenter, 48.5; and unskilled workman 44.4. (1, p. 192) Public administrators, individually, differ considerably in their scores on the OL scale, the range being from 32 to 82, 94 per cent scoring 50 and higher, 81 per cent scoring 55 and higher, 58 per cent scoring 60 and higher and 30 per cent scoring 65 and higher.

On these two interest scales—public administrator and OL—public administrators score very much alike; only a small minority scoring appreciably lower than the remainder.

On the other hand, public administrators as a whole score surprisingly

low on 34 other interest scales. (See Table 2.) In terms of ratings they obtain no A or B+ ratings and only three B ratings,[1] e.i., 39 on personnel manager interest, 38 on production manager interest and 35 lawyer interest. Is this poor showing on 34 occupational interests because public administrators do not have such interests or because they differ so much among themselves that high scores of some administrators are offset by low scores of other administrators?

Per Cent of A Ratings. Table 1 records the number of A ratings that would be received if there were 100 men in each of the 16 sub-groups, in other words, the data are percentages. Thirty-three per cent of all the public administrators represented here received an A rating in personnel manager interest, 30 per cent in production manager, 20 per cent in engineer interest, 18 per cent in both accountant and city school superintendent interest and at the other extreme, 1 per cent in dentist, 2 per cent in carpenter, 3 per cent in musician, artist and architect and 4 per cent in mathematician interest. The average for all 34 occupational interests is 10.4 per cent, which means that on the average a public administrator obtains three and a half A ratings on 34 scales.

But Table 1 discloses that there is great variation in A ratings not only with respect to the 34 occupational interests but also with respect to the 16 sub-groups of public administrators.

A criterion group has approximately 70 per cent A ratings on its own scale. There are five instances in Table 1 where sub-groups obtain approximately that percentage of A ratings on a scale. City school superintendents have 74 per cent of A ratings on the school superintendent interest scale, engineers have 73 per cent A ratings on the production manager scale, forest service supervisors have 72 per cent A ratings on the forest service scale and chemist-physicists have 69 per cent A ratings and engineers 62 per cent on the engineer interest scale. The city school superintendents were all included in the criterion group and 41 of the 50 forest service supervisors were included in that criterion group. On this basis the high percentages in these two cases may be challenged although 67 per cent of the nine forest supervisors not included in the criterion group also obtained A ratings. But this cannot be said of the engineers who scored as high as the original criterion group on the production manager scale nor of the chemist-physicists and engineers who scored approximately as high as the original criterion group on the engineer scale.

[1] Scores of 45 and higher are rated A; 40 to 44, B+; 35–39, B; 30–34, B−; 25–29, C+; and 24 and below, C. A rating of A means one has the interests of the occupation, a rating of C means one does not have those interests, a rating of B is intermediate implying one probably has those interests.

Table 1

"A" Ratings on 34 Occupational Interests of 16 Sub-groups of Public Administrators. Data are Percentages

Table 1

"A" Ratings on 34 Occupational Interests of 16 Sub-groups of Public Administrators. Data are Percetanges

Group	Scale	Personnel	Welfare	Taxation	Comptr. Finance	Recreation	Statistician	Public Health	Engineer	Chemist-Physicist	Prison Warden	Hosp. Supt.	City Mgr.	For. Serv. Supr.	City Sch. Supt.	For. Serv. Admin.	Fed. Admin.	Average
	N =	61	50	15	24	17	27	25	26	14	16	45	41	50	50	38	53	552
I	Artist	2		7	4	12	4	8							1	3	2	2.7
	Psychologist	2	8	7	8		30	12	8	23	6	2	2		1	5	2	7.3
	Architect		4	7		6	4		12	15						3	2	3.3
	Physician	2	2	7	4	6	11	32	4	15	6	2		8	2	8	4	7.1
	Dentist	2						4						2			2	0.6
II	Mathematician		2		4		19	4	12	23			2			3		4.3
	Engineer	10	12	7	8		37	12	62	69	13	4	29	14	4	22	9	19.5
	Chemist	8	2	7	12		30	16	38	54	6	4	10	4		14	8	13.3
III	Production Mgr.	36	33	27	29	24	30	8	73	54	6	33	32	22	7	35	25	29.6
IV	Farmer		10			6	11	16	12	23	6	2	12	32	3	16	13	10.1
	Carpenter	2	4				4		8	8			2	8				2.3
	Printer	5			4	6	4	8	8	15	6		2	2				3.8
	Math. Sci. Teacher	12	6	7	9	35	19	20	20	23	6	9	12	4	17	5	6	13.1
	Policeman	2	6	7	5	6	4		8		6	2	10	10	1	3	6	4.8
	Forest Service	5	6		4	6	7	12	12		13	4	12	72	3	30	13	12.4
V	Y. Physical Dir.	5	16	7	8	35	4				6	13	2	4	14	3		7.3
	Personnel Mgr.	51	48	47	46	47	37	4	19	8	38	49	25	14	25	35	34	32.9
	Y.M.C.A. Secy.	5	23	20	4	41	4	4			19	13	2	8	20	3	6	10.8
	Soc. Sci. Teacher	22	29	7	20	40	11	8			19	11	7	14	29	11	8	14.8
	City Sch. Supt.	22	35	13	8	41	15	12			13	18	10	4	74	8	19	18.3
	Minister	5	16	13		41		8			13	4		2	15	5	2	7.8
VI	Musician	2	2	7		18	4	4							3			2.5
VII	C.P.A.	7	8	27	21	6	11		8			4	20		2		6	7.7
VIII	Accountant	19	18	47	46	29	26		8	15	6	22	27	10	3	8	11	18.4
	Office	12	22	40	46	24	11	4	4	8	6	24	15	14	7	14	4	15.9
	Purchasing Agent	15	12	20	17	6	7	4	12	15		22	22	8		8	9	11.1
	Banker	3	10	7	24	12		4			6	11	7	6	6	14	6	7.3
IX	Sales Manager	14	6	20	16	12	4	8				18	7	14	4		13	8.5
	Real Estate SLM.	15	10		8	12		12			6	7	12	14	3		4	6.4
	Life Insur. SLM.	13	10	7	8	12	4	12				13	5	6	11		4	6.6
X	Advertising Man	17	18	7	9	24	7	12	4		13	7	2	4	9	3	9	9.1
	Lawyer	27	22	33	21	35	11	24	4		13	7	8	16	22	11	28	17.6
	Author-Journal.	7	16	7	4	12	7	16	4			2			7	8	4	5.9
XI	President	5	12	7	12		4	12	27	23	6	18	20	10	5	8	11	11.3
	Average	10	13	12	12	16	11	9	11	12	7	10	9	10	9	9	8	10.4
	Pub. Admin.	78	80	87	75	71	71	40	54	57	56	67	54	38	—	73	81	65.4

Each one of the sixteen sub-groups obtains at least 30 per cent A ratings on one occupational interest but they differ greatly with respect to the occupational interests on which they score that high.

The sixteen sub-groups of public administrators may be classified on the basis of the occupational interests on which they receive 30 per cent or more A ratings as follows:

1. Engineers and chemist-physicist administrators have 38 to 73 per cent A ratings on engineer, chemist and production manager interests.

2. Statisticians also have 30 per cent or more A ratings on these three occupational interests but they have an average of only 32 per cent such ratings while engineers and chemist-physicists have 58 per cent. Furthermore they have such A ratings on psychologist and personnel interests and engineers and chemist-physicists have only half that percentage. On still other grounds to be given below it seems best to group statisticians separately from the other two.

3. City managers have 32 per cent A ratings on production manager and 29 per cent on engineer interest. Additional data given below place it in the same general area as 1 and 2.

4. Public health officials have 32 per cent A ratings on physician interest.

5. Forest service supervisors have 72 per cent A ratings on forest service interest and 32 per cent on farmer interest.

The above six sub-groups differ from the average of all in having stronger scientific, production or mechanical interests and weaker social service, public accounting, clerical or sales interests. The next four sub-groups differ from the average in reverse manner.

6. Taxation and comptroller-finance administrators have 30 per cent or more A ratings on accountant, office and personnel interests.

7. City school superintendents have 74 per cent A ratings on school superintendent interest.

8. Recreation administrators have 35 to 41 per cent A ratings on all six occupational interests in Group V [2] and also on lawyer and mathematics-science teacher interests. This is the only sub-group of public administrators that scores high on Y.M.C.A. physical director, Y.M.C.A. secretary, social science teacher, minister and mathematics-science teacher interests.

9. The remaining six sub-groups, i.e., personnel, welfare, prison warden, hospital superintendent, federal and forest service administrators have 30 per cent or more A ratings on personnel interest and all but

[2] For the sake of convenience the 34 occupational interests are classified under eleven groups, see Table 1. Groups are composed of occupational interests which correlate .60 or higher with one another.

prison warden has such A ratings on production manager. In addition, the welfare sub-group has 35 per cent A ratings on school superintendent interest and forest service administrator has 30 per cent A rating on forest service interest.

The above gives a rough classification of public administrators.

The data in Table 1 indicate that certain sub-groups are much alike and certain sub-groups differ greatly from other sub-groups in their interests. References to the table will show, for example, that recreation and chemist-physicists are largely reversed in their interests.

The differences in interests between sub-groups explains why public administrators on the average obtain relatively few high interest scores on the 34 occupational scales—the high scores of some public administrators are offset by low scores of other administrators.

Mean Interest Scores. Per cent of A ratings gives a good idea of how many men score high in a given interest. Mean scores present a slightly different picture taking into account the low scores as well as the high scores. In this case the conclusions drawn from the two sets of data agree very closely.

Mean scores of the 16 sub-groups of public administrators on 34 occupational interests are given in Table 2. The table contains a wealth of information regarding the interests of these public administrators. It shows that they all score relatively high on production manager interest and very low on carpenter interest. It shows that public health administrators score highest of all on physician interest and comptrollers the lowest. It shows that engineers and chemist-physicists score high and recreation administrators very low on engineer, chemist and production manager interests and that the reverse is the case with respect to Y.M.C.A. physical director, Y.M.C.A. secretary, social science teacher, city school superintendent and minister interests. It shows that sub-groups of public administrators are much alike on some scales, such as printer and policemen interests, and that they differ among themselves very greatly on other scales, such as engineer and forest service interests.

How Much Do Public Administrators Differ?

It is important to know not only how each sub-group differs from other sub-groups but also how much. Because of the exceedingly large number of inter-relationships between 16 sub-groups expressed on 34 interest scales it has seemed desirable to present our findings in three different ways. First, all statistically significant differences between each sub-group and the remaining 15 sub-groups are considered. Second, all statistically significant differences between each sub-group and the average of all are pointed out. Third, all differences between each sub-

Table 2

Mean Scores of 16 Sub-groups of Public Administrators on 34 Occupational Interests

	Chem.-Physicist	Engineer	Public Health	F. S. Superv.	Stat.	F. S. Admin.	City Mgr.	Prison Warden	Federal	Hosp. Supt.	Personnel	Welfare	Taxation	Comptroller	School Supt.	Recreation	Average
Artist	27	24	27	21	24	21	17	21	17	17	20	21	19	14	20	21	21
Psychologist	31	24	25	11	32	21	17	19	20	18	23	24	21	15	22	21	21
Architect	32	32	27	22	28	24	24	22	20	19	21	23	23	16	21	21	23
Physician	36	33	40	27	32	29	24	27	22	27	23	25	22	20	27	25	27
Dentist	28	26	32	24	23	21	20	22	19	24	18	19	16	16	20	20	22
Mathematician	36	30	28	15	32	23	25	18	19	17	19	19	21	17	25	18	23
Engineer	48	49	33	33	38	35	38	28	31	28	28	27	28	28	25	23	33
Chemist	44	43	33	24	37	30	29	22	27	24	26	25	23	22	21	20	28
Production Mgr.	42	49	33	40	39	40	42	37	41	40	38	39	37	41	31	34	39
Farmer	38	33	34	41	32	37	33	31	34	29	28	28	26	28	29	30	32
Carpenter	25	25	18	25	21	20	21	17	16	16	15	17	14	13	11	18	18
Printer	33	30	26	27	31	27	27	29	28	25	29	30	29	25	23	31	28
Math. Sci. Teacher	36	33	31	30	34	31	32	29	33	31	32	32	31	31	34	35	32
Policeman	29	28	29	33	28	30	30	32	30	29	29	30	28	30	24	32	29
Forest Service	29	31	30	49	29	40	29	30	33	27	28	29	26	25	22	31	31
Y. Physical Dir.	22	20	28	28	25	27	23	28	28	31	29	32	27	27	31	40	28
Personnel Mgr.	31	35	30	34	39	39	35	41	43	42	47	46	42	43	38	44	39
Y.M.C.A. Secy.	20	19	24	25	25	27	24	29	30	33	30	35	32	30	35	41	29
Social Sci. Teacher	22	21	27	28	30	30	27	35	33	33	37	37	35	35	39	44	32
City School Supt.	24	24	33	27	31	32	29	34	33	34	36	37	35	33	50	42	33
Minister	19	16	23	19	24	22	17	25	22	26	24	31	26	21	33	35	24
Musician	19	19	23	18	25	20	17	21	19	18	23	25	23	17	22	29	21
C.P.A.	27	19	26	20	36	26	32	29	29	27	32	30	38	38	28	28	29
Accountant	28	32	23	28	35	30	38	32	35	35	34	34	42	44	29	33	33
Office	26	29	26	31	30	31	35	33	33	37	34	34	38	42	31	34	33
Purchasing Agt.	29	35	27	34	29	31	37	31	34	36	33	29	34	38	26	29	32
Banker	25	25	31	33	28	31	33	33	31	33	30	30	34	37	32	31	31
Sales Manager	24	31	30	33	26	30	32	31	34	35	34	33	33	36	30	32	32
Real Estate Slsmn.	26	29	32	36	28	31	33	34	33	33	34	33	32	35	32	34	32
Life Insur. Slsmn.	19	23	29	36	24	31	33	31	33	33	32	32	30	32	32	33	30
Advertising Man	28	29	32	28	31	29	27	33	31	30	34	34	32	30	30	33	31
Lawyer	31	30	35	34	35	33	31	36	32	30	37	36	36	34	38	36	34
Author-Journalist	32	30	34	28	33	30	27	32	28	27	31	31	29	27	31	31	30
President	33	40	33	33	34	34	34	34	34	40	34	34	34	34	34	34	34

group and the reaminder that amount to over-lapping between distributions of 50 per cent and less are discussed.

Statistically Significant Differences. The statistical procedure employed here is as follows. The difference in score between each pair of sub-groups on each scale was noted; the critical ratio of that difference determined, also the per cent of over-lapping between the two distributions. For example, consider the scores of sub-groups on the artist interest scale as given in Table 2. Comparisons between public health and five other sub-groups are as follows:

Sub-Groups	Difference in Score	Critical Ratio	Per Cent Overlapping[3]
Statistician	2	.7	92
Personnel	6	2.5	76
Taxation	7	2.0	74
City Manager	8.9	3.7	63
Comptroller	11.4	3.9	58

From 334 such calculations the following relationships were established:

Difference in Standard Scores between Two Means	Critical Ratio of Difference	Per Cent of Overlapping
5.1	2	80
8.0	3	70
11.3	4	60
14.4	5	50

The correlation between critical ratio and per cent of over-lapping is −.988. Since critical ratio is affected by size of population and per cent of over-lapping is not, the above relationships hold for only the populations that are contained in these sub-groups. To escape the labor of calculating critical ratios and percentages of overlapping for the several thousand relationships in Table 2 we have assumed that a difference of 8 gives a critical ratio of 3 and a difference of 14 gives a critical ratio of 5 or an overlapping of 50 per cent. This is not strictly true for a difference of less than 8 between two of the larger sub-groups gives a critical ratio of 3 and a difference of 8 between two of the smaller sub-groups does not give a critical ratio of quite 3. But the exceptions are surprisingly few among the 334 cases. On this basis we shall consider two sub-groups as

[3] The formula used is: $\frac{\text{Mean}_1 - \text{Mean}_2}{\frac{1}{2}(\text{Sigma}_1 + \text{Sigma}_2)}$. Reference to Tilton's table gives the per cent of overlapping for the obtained quotient (5).

different in their interests if their scores differ by as much as 8, which gives approximately a critical ratio of 3 and 70 per cent overlapping in the distributions of the two sets of scores.

The number of such differences is recorded in Table 3 for each of the 120 pairs of sub-groups. There are all told 870 statistically significant differences, or an average of 7.3 differences per pair. The 16 sub-groups differ greatly in this respect. Prison warden, federal and personnel average 4.1 to 4.7 and engineer and chemist-physicist, 12.1 to 13.5 differences.

Table 3

Number of Significant Differences in Scores on 34 Occupational Interest Scales between Pairs of 16 Sub-Groups of Public Administrators

	Fed.	P	FA	HS	W	Tx.	CM	St.	FS	SS	PH	CF	R	E	CP	Ave.
Prison warden	0	0	2	0	0	2	3	4	5	4	7	3	5	13	14	4.1
Federal	—	0	0	0	1	2	1	5	5	5	6	4	7	14	15	4.3
Personnel		—	3	0	0	1	3	7	8	3	7	3	3	16	17	4.7
Forest admin.			—	2	4	4	2	4	3	10	4	9	9	9	10	5.0
Hospital supt.				—	0	2	5	9	6	5	11	3	7	15	19	5.6
Welfare					—	2	8	6	10	2	10	8	1	14	18	5.6
Taxation						—	3	6	9	3	10	0	5	18	20	5.8
City manager							—	5	5	11	8	6	11	6	13	6.0
Statistician								—	7	12	5	15	11	6	4	7.1
Forest supervisor									—	13	6	8	13	11	11	8.0
School supt.										—	9	8	5	16	17	8.2
Public health											—	15	13	8	6	8.3
Comptroller-finance												—	10	19	19	8.7
Recreation													—	15	19	8.9
Engineer														—	1	12.1
Chemist-physicist															—	13.5

The first three sub-groups in the table, namely, prison warden, federal and personnel, do not differ significantly from one another on any of the 34 occupational scales and they differ least of all from the remaining thirteen sub-groups. Nevertheless they differ from many of the thirteen in a number of interests.

The next four sub-groups, namely forest service administrator, hospital superintendent, welfare and taxation, differ very little from the first three (average of 0.9 occupational interests per pair) and they differ relatively little among themselves (average of 2.3 interests per pair). But they differ much more than this from most of the remaining sub-groups.

The next seven sub-groups differ appreciably among themselves and with the preceding seven sub-groups. There is one exception to this

generalization. Comptroller-finance and taxation administrators do not differ significantly from one another on any of the 34 occupational interests; the former, however, differ half again as much from the other sub-groups as does the latter.

The last two sub-groups, engineer and chemist-physicist, differ significantly from the other fourteen on four to twenty occupational interests but they differ from each other on only one interest, that of C.P.A.

Eleven pairs of administrators do not differ significantly on any occupational interest, i.e.,

Prison warden—Federal administrator
Prison warden—Personnel man
Prison warden—Welfare official
Prison warden—Hospital superintendent
Federal administrator—Personnel
Federal administrator—Forest service administrator
Federal administrator—Hospital superintendent
Personnel—Welfare
Personnel—Hospital superintendent
Welfare—Hospital superintendent
Taxation—Comptroller-finance

The following thirteen pairs differ significantly on only one or two occupational scales.

Prison warden—Forest service administrator
Prison warden—Taxation
Federal—Walfere
Federal—Taxation
Federal—City manager
Personnel—Taxation
Forest service administrator—Hospital superintendent
Forest service administrator—City manager
Hospital superintendent—Taxation
Welfare—Taxation
Welfare—School superintendent
Welfare—Recreation
Engineer—Chemist-physicist

The above 24 pairs may be considered to have very similar interests so far as these data go. The remaining 96 pairs of administrator sub-groups range all the way from a few significant differences to many such differences some of which are large in amount.

Sub-Groups Compared with Median of All Groups

It is interesting to know how various sub-groups of public administrators differ from one another. It is also worthwhile to know how each differs from all. A sub-group may be considered to differ from public administrators in general if its mean score is 8 or more higher or lower than the median score of the sixteen sub-groups.[4] In such a case the sub-group would overlap by 70 per cent or less the distributions of at least half of the remaining sub-groups.

The data in Table 2 indicate that the four sub-groups of prison warden, federal, personnel and hospital superintendent do not differ by as much as 8 scores from the median of all public administrators on any of the 34 occupational interests.[5] On this basis they may be said not to differ from public administrators in general. The data in Tables 2 and 3 make clear, however, that they do differ significantly from most sub-groups in several interests. But such differences are largely caused by the other sub-groups deviating considerably from the median of all.

Forest service administrators, welfare and city managers differ by 8 scores or more from the median of all on only one occupational interest each. Forest service administrators score higher than the median on forest service interest, welfare score higher on minister interest and city manager on engineer interest.

A diagrammatic presentation of the relationship between the 16 sub-groups is given in Figure 1. Within the inner circle are the four sub-groups which do not differ on any scale from the median of all. On the second circle are the three sub-groups which differ on one scale from the median. On the third circle are city school superintendent and taxation which differ on two scales from the median. Comptroller-finance which differs on three scales is placed between the third and fourth circles. The remaining six sub-groups which differ on 5 to 15 scales are placed on the fourth, or outer circle. As nearly as is possible to represent the relationships in two dimensional space the 16 sub-groups are assigned positions in the direction of the major interest which distinguishes them from the median of all sub-groups. These major interests are recorded outside the four circles.

Taxation administrators differ significantly from the median of all on two occupational interests, namely C.P.A. and accountant. Comptroller-finance administrators differ on the above two interests plus that of

[4] The median scores of the 16 sub-groups differ on the average by 0.67 from the average of the 620 public administrators, given in Table 2. In only two cases do the median and average differ as much as 2.0.

[5] Prison warden does not differ from the median on any scale by as much as 4, federal differs on only two scales by that amount, personnel differs by 4 on one scale and 7 on another, hospital superintendent differs by 4 on three scales and by 4.5 on two scales.

office worker. In both these sub-groups there is a general trend toward higher scores on all Group VIII interests and toward lower than median scores in Groups I, II and IV, which sets these two sub-groups off from the remainder.

School superintendents score significantly higher than the median of all sub-groups on school superintendent and minister interests and lower on production manager interest.

The remaining six sub-groups score significantly different from the median of all on 5 to 15 occupational interests.

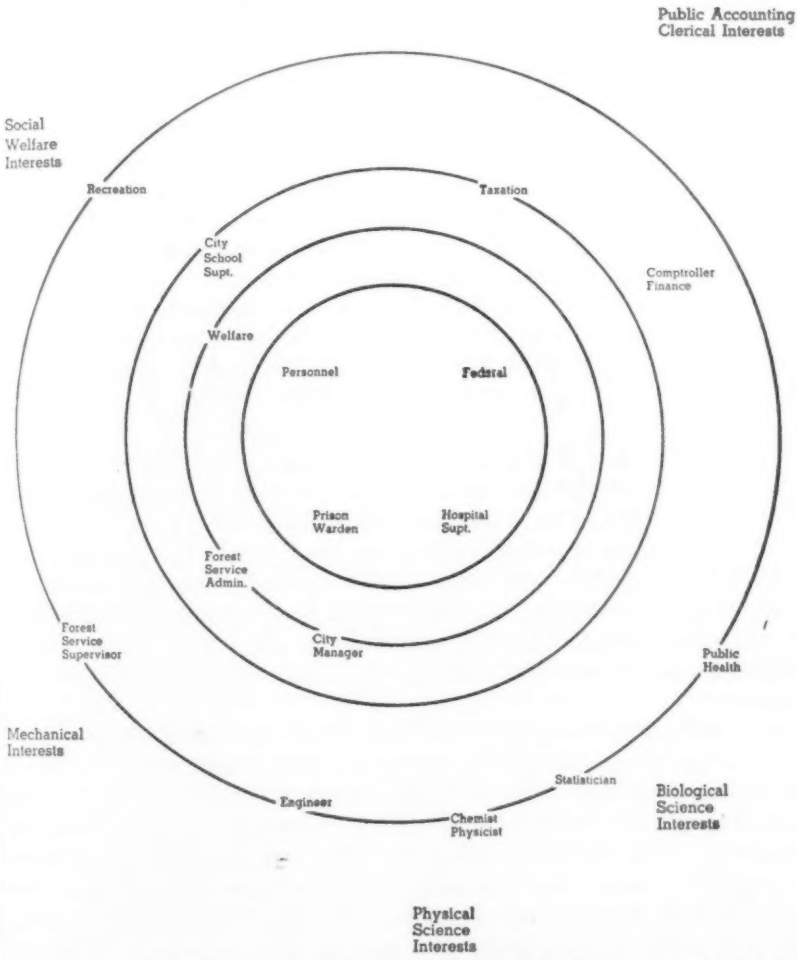


FIG. 1. Diagram showing how 16 sub-groups of public administrators differ from public administrators as a whole.

Forest service supervisor scores higher on forest service, farmer and carpenter interests and lower on psychologist and C.P.A. interests. The striking thing about the interests of this sub-group is that they are in Group IV. No other sub-group scores as high in farmer and forest service interests, although forest service administrators, scoring 9 below the supervisors, on the forest service scale, still score significantly above the median of all on that scale. Forest service supervisors share with engineers and chemist-physicists the distinction of scoring 7.5 above the median on carpenter interest.

Statisticians are a heterogeneous group as far as education and interests are concerned and we doubt whether such a grouping is significant. This is all the more surprising since nearly all of our statisticians are department heads in the Census Bureau. As it is, they score higher than the median of all groups on mathematician, chemist, psychologist, engineer and C.P.A. interests and lower on life insurance interest. They are similar to engineers and chemist-physicist in having stronger than average scientific interests and weaker sales interests but they differ from these two groups in having stronger than average public accounting interest and weaker production manager interests. They differ from every other group on four to fifteen different interests, averaging 7.1 interests. They differ least—on only 4 interests—from prison warden, forest service administrator and chemist-physicist; they differ most—on 15 interests—from comptroller-finance, despite the fact that both score markedly higher than the median on C.P.A. interest.

The public health sub-group scores significantly higher than the median of all sub-groups on physician, dentist, mathematician and chemist interests and lower on personnel and accounting interests. The trend here is higher scientific interests and lower production management, personnel management and office activity interests than the remainder. This sub-group differs on the average from the remaining 15 sub-groups by 8.3 occupational interests, differing the least from forest service administrators (4 interests) and the most from comptroller-finance (15 interests).

The recreation sub-group scores close to the median of all sub-groups on 28 of the 34 occupational interests but differs significantly on five of the six social interests (Group V) and on musician interest. The strong interest in these two directions sets off the recreation sub-group from the remainder. The recreation sub-group differs significantly from welfare in only one interest, from personnel in only three interests and from prison warden, taxation and school superintendent in five interests each. One explanation for the general agreement in interests between recreation and these five sub-groups is that the five average slightly above the median of all public administrators on social and musical interests.

Chemist-physicist and engineer sub-groups are the most conspicuous examples of sub-group scoring 8 or more scores above or below the median of all sub-groups. The former so differs on 15 interest scales, the latter on 12 scales. They score higher than the median in scientific, production manager and president interests (Groups I, II, III and XI) and lower in public accounting, selling and social interests (Groups VII, IX and V). They differ significantly on the average from other public administrators on 12 to 13 scales (Table 3).

Because they differ from each other on only one occupational interest, namely, C.P.A., it can be argued that the two sub-groups should be combined. This is a tenable procedure. The fact that one group is composed of engineers engaged in engineering administration for the Tennessee Valley Authority, U. S. Forest Service, Geodetic Survey and Department of Agriculture, also as state highway engineers and commissioners of public works and the other group directs the activities of chemists and physicists in the Weather Brueau, the Bureau of Standards and the Department of Agriculture warrants two sub-groups for they are engaged in two fairly distinct activities.

Overlapping of 50 Per Cent or More

The writer recalls no discussion of the question, When are two groups different? It is customary to accept two averages as different if they equal or exceed three times the standard deviation of the difference, which as far as these data go, means an overlapping of 70 per cent or less. The writer is not at all sure that two groups would always be viewed as different when their mean scores are significantly different. It is entirely possible that opinion would differ somewhat according to the attribute measured.

So that no one can challenge our standard of difference between two groups we present here differences between sub-groups of public administrators where their scores on an interest scale overlap fifty per cent or less. As explained above such differences correspond to differences of 14 in score. Because most of the differences of 14 and more are larger than 14 the overlapping in the tables to follow is less than fifty per cent.[6]

We have recorded every difference in score of 14 or more between the means of each sub-group and every other sub-group basing the calculations on mean scores in Table 2. The resulting 200 differences are recorded in Table 4. The notation in the cells of the table indicate the

[6] From the 15 differences in the 16 upper left hand cells and the 19 differences in the 4 lower left hand cells of Table 4 we find the differences in scores range from 13 to 27 (average of 18.6), the critical ratios range from 3.5 to 13.5 (average of 6.3) and the per cent of overlapping ranges from 55 to 14 (average of 38 per cent). The difference of 13 has an overlapping of 44 per cent. The critical ratio of 3.5 accompanies a difference of 17 and has an overlapping of 52 per cent.

Table 4

Differences in Score of 14 and More on Occupational Interests of 16 Sub-groups of Public Administrators

	Chem.-Phys.	Eng.	Public Health	F. S. Superv.	Stat.	F. S. Admin.	City Mgr.	Prison Warden	Federal
Chemist-Physicist	—	0	E	Ps. Ma. E. Ch.	0	Ch.	Ps. Ch.	Ma. E. Ch.	D. Ma. E. Ch.
Engineer	0	—	E. Pm.	Ma. E. Ch.		E	Ch.	E. Ch.	E. Ch.
Public Health			—	Ps.	0	0	D	0	D
Forest Serv. Supervisor	Fs. LI	FS	FS	—	FS	0	FS	FS	FS
Statistician	0	CPA	0	Ps. Ma. CPA	—	0	Ps.	Ma. Ch.	0
Forest Serv. Admin.			0	0	0	—	0	0	0
City Manager	LI		Ac			0	—	0	0
Prison Warden		ST	0			0	0	—	0
Federal Admin.	LI				0	0	0	0	—
Hospital Supt.	LI	YS	0			0	0	0	0
Personnel	P. ST.	ST.	P		0	0	0	0	0
Welfare	P. YS. ST	YS. ST. M	P		0	0	M	0	0
Taxation	Ac	ST. CPA	Ac	CPA. Ac			0	0	0
Comptroller-Finance	Ac. Of.	ST. CPA	Ac. Of.	CPA. Ac		Ac	0	0	0
City School Supt.	YS. ST. SS. M.	YS. ST. SS. M.	SS.	SS. M	SS	SS	SS. M	SS	SS
Recreation	YP. YS. ST. SS. M. LI.	YP. YS. ST. SS. M.	P. YS. ST.	YS. ST. SS. M.	YP. YS. ST.	YS. ST.	YP. YS. ST. M.	0	0

Recreation	YP. YS. ST. SS. M. LI.	YP. YS. ST. SS. M.	P. YS. ST.	YS. ST. SS. M.	YP. YS. ST.	YS. ST.	YP. YS. ST. M.	0	0

Table 4—*Continued*

	Hosp. Supt.	Person.	Welfare	Taxation	Comptroller	School Supt.	Recreation
Chemist-Physicist	Ma. E. Ch.	Ma. E. Ch.	Ma. E. Ch.	D. Ma. E. Ch.	Ps. Ar. D. Ma. E. Ch.	E. Ch. Ca.	Ma. E Ch.
Engineer	E. Ch.	E. Ch.	E. Ch.	E	Ar. E. Ch.	E. Ch. Pm. Ca.	E. Ch. Pm
Public Health	0	D. Dt.	D	D. Dt.	D. Dt.		D
Forest Serv. Supervisor	FS	FS	FS	F. FS.	FS	Ca. FS	FS
Statistician	Ps. Ma.	0	0	Ch	Ps. Ma. Ch.	Ch.	Ma. E Ch.
Forest Serv. Admin.	0	0	0	FS	FS	FS	
City Manager	0	0		0	0		E
Prison Warden	0	0	0	0	0		0
Federal Admin.	0	0	0	0	0		0
Hospital Supt.	—	0	0	0	0		0
Personnel	0	—	0	0	0		0
Welfare	0	0	—	0	0	0	0
Taxation	0	0	0	—	0		0
Comptroller-Finance	0	0	0	0	—	Ac	
City School Supt.	SS	SS	0	SS	SS	—	0
Recreation	0	0	0	0	M	0	—

Abbreviations of Interest Scales

Ac	Accountant
Ar	Architect
Ca	Carpenter
Ch	Chemist
CPA	Certified Public Account.
D	Doctor
Dt.	Dentist
E	Engineer
F	Farmer
FS	Forest Service
LI	Life Ins. Salesman
M	Minister
Ma	Mathematician
Of	Office man
P	Personnel
PM	Production Manager
Ps	Psychologist
SS	City Sch. Supt.
ST	Soc. Sci. Teacher
YP	Y. Phys. Director
YS	Y.M.C.A. Secretary

occupational interests on which the sub-group at the left scores higher than the sub-group at the top of the column. To ascertain all the differences in interests between two sub-groups it is necessary to look in two places, i.e., opposite the name of both sub-groups and beneath the name of the other sub-group. For example, engineers score 14 or more higher than forest supervisors on mathematician, engineer and chemist interests and score 14 or more below forest supervisor on forest service interest. A "0" in a cell means that there are no differences of 14 or more between the two sub-groups on any one of the 34 occupational interests. As an example, see the cell opposite chemist-physicist and below engineer, also the corresponding cell opposite engineer and below chemist-physicist. When no figure appears in a cell it means that there are one or more differences between the two sub-groups recorded in the other cell common to the two. For example, see opposite public health and below chemist-physicist and then note opposite chemist-physicist and below public health.

The 16 sub-groups have been arranged in rank-order according as they have stronger interests than the majority in scientific, production and mechanical interests (Groups I, II, III and IV) and weaker than the majority in social, public accounting, clerical and sales interests (Groups V, VII, VIII and IX). In passing it is worthy of note that no two sub-groups of public administrators differ by scores of 14 in musical, linguistic[7] or president interests (Groups VI, X and XI).

On the basis of Table 4 the 16 sub-groups may be divided into three rather distinct classes: (1) those who differ noticeably from the average in scientific, production or mechanical interest, (2) those who differ noticeably from the average in social, accounting, clerical or sales interests and (2) those who fall in between and differ very little from the average in either respect. A summary of the differences in interests between the three classes is given in Table 5.

The first five sub-groups, namely, chemist-physicist, engineer, public health, forest supervisor and statistician score higher than other sub-groups on 113 measures and all but three of these measures are in scientific, production or mechanical interests.

The last four sub-groups, namely, taxation, comptroller-finance, school superintendent and recreation, score higher than other sub-groups on 65 measures, all of which are social, accounting, clerical or sales interests.

The middle seven sub-groups, namely, forest service administrator, city manager, prison warden, federal administrator, hospital superintend-

[7] This would not be the case if there were included here public administrators engaged primarily in legal or publicity work. The few cases we have of such men score high on Group X.

ent, personnel and welfare score higher than other sub-groups on 22 measures, 4 of which are scientific, production of mechanical interests and 18 in the opposite direction.

Members of each of these three classes are much alike. The seven sub-groups in Class 2 do not differ from one another except on one scale. The four sub-groups in Class 3 differ among themselves on only 4 measures. And the five sub-groups in Class 1 differ from one another on 20 measures. Difference, as used here, it must be remembered, constitutes overlapping between distributions of 50 per cent or less.

Classes 1 and 3 differ markedly from each other. Thus, Class 1 scores higher than Class 3 on 45 measures and Class 3 scores higher than

Table 5

Summary of Table 4
Number of Differences in Interests between Sub-Groups according to Membership in Three Classes

	Class 1	Class 2	Class 3	Total
Class 1	20	48	45	113
Class 2	17	1	4	22
Class 3	47	14	4	65
Total	84	63	53	200

Per pair of sub-groups

	Class I	Class II	Class III	Total
Class 1	1.00	1.14	2.25	1.51
Class 2	.49	.02	.14	.21
Class 3	2.35	.50	.33	1.08
Total	1.12	.60	.88	.83

Class 1 on 47 measures, a total of 92 of the 200 differences. A fairer comparison is in terms of differences per pair, see Table 5. On this basis, Class 1 scores higher than Class 3 2.25 times per pair and Class 3 scores higher than Class 1 2.35 times per pair, in contrast to .83 per pair for the 240 pairs. Class 2 is much more similar to Class 3 than to Class 1 for Class 2 differs from Class 3 only 18 times (.32 per pair) while Class 2 differs from Class 1 65 times (.93 per pair).

The five sub-groups constituting Class 1 are not entirely alike in their interests. In Table 6 are recorded the number of times the five sub-groups score higher than other sub-groups and in Table 7 the number of times they score lower than other sub-groups. On this basis public health officials score high primarily in biological scientific interests (Group I), forest service supervisors in mechanical interests (Group IV), and engineers in physical science interests (Group II), whereas chemist-

Table 6

Number of Times the 9 Sub-Groups in Classes 1 and 3 Score Higher than the Remaining 15 Sub-Groups. Scores are Distributed by Occupational Groups *

Occupational Groups	Class 1					Class 3				
	Chem-Phys.	Eng.	Statistician	Pub. H.	Forest Supv.	Taxation	Compt. Fin.	Welfare	Sch. Supt.	Recreation
I	7	1	4	11						
II	32	23	11							
III		3								
IV	1	1			16					
V						1	1	8	21	27
VII			2			2	2			
VIII						3	7			
IX					1					1

* Differences are 14 or more, equivalent to differences in overlapping of 50 per cent or less.

The welfare sub-group of Class 2 is included here to show its relationship to the four sub-groups of Class 3.

Table 7

Number of Times the 9 Sub-Groups in Classes 1 and 3 Score Lower than the Remaining 15 Sub-Groups. Scores are Distributed by Occupational Groups *

Occupational Groups	Class 1					Class 3				
	Chem-Phys.	Eng.	Statistician	Pub. H.	Forest Supv.	Taxation	Compt. Fin.	Welfare	Sch. Supt.	Recreation
I					3	3	7	1		1
II				2	7	5	7	5	5	9
III				1					1	1
IV	1		1	1		3	2	1	5	1
V	14	18	4	6	6	1	2			
VII		3			3					
VIII	3			4	2				1	
IX	5									

* See footnote to Table 6.

physicists and statisticians score high primarily in Group II and to a lesser degree in Group I. Chemist-physicists and engineers score low in social interests (Group V), much lower than public health, forest service supervisors and statisticians. Forest service supervisors, unlike the other four sub-groups, not only do not score higher than other sub-groups in scientific interests but actually score lower than a few sub-groups.

In Class 3, taxation and comptroller-finance differ from school superintendent and recreation in that the former score high in public accounting

and office activities (Groups VII and VIII) and the latter score high in social welfare activities (Group V). All four sub-groups score low in physical, scientific and mechanical activities (Groups II and IV). The welfare sub-group could be assigned to Class 3 just about as well as to Class 2. To show its similarity to school superintendent and recreation it is included in Tables 6 and 7 as though it were a member of Class 3.

Summary

The group of 552 public administrators are drawn from a wide variety of public activities, earning from three to over ten thousand dollars a year, with an average income of six thousand dollars. Three-fourths of them are college graduates.

They have been divided into sixteen sub-groups on the basis of function pursued.

Their interests have been recorded on the Vocational Interest Blank, which has been scored on 34 occupational interest scales. The basic data are 34 mean scores for each of 16 sub-groups.

Public administrators, as a whole, obtain no A or B+ ratings on the 34 occupational scales. Their three B ratings are on personnel (score of 39), production manager (38) and lawyer interest (35). The fact that they score highest on personnel and production manager interests is suggestive of their two major functions, i.e., handling people and getting certain work done.

The reason they average so low on the 34 occupational interests is that they differ in their interests, the high scores of some administrators being offset by low scores of other administrators.

Differences between sub-groups are given in terms of overlapping of 70 per cent or less (critical ratio of 3 or more) and overlapping of 50 per cent or less. On the former basis there are 870 significant differences between the 120 pairs of sub-groups, but 11 pairs of sub-groups have no such difference on any of the 34 occupational interest scales. On the latter basis there are 200 differences between pairs with 52 pairs not differing on any scale.

Sub-groups differ from one another on nearly all the 34 occupational scales but less on the scales in Groups VI, IX, X and XI which are concerned with musician, sales, linguistic, and president interests. The few cases we have of administrators engaged primarily in legal and publicity work indicate that if such men had been included there would have been noticable differences between them and the other sub-groups on Group X interests.

A classification of the 16 sub-groups, as far as interests are concerned, is given in terms of how each sub-group differs from the median of the 16 sub-groups. Essentially the same classification emerges if the sub-groups are grouped according to their highest occupational interest scores.

The five sub-groups of hospital superintendent, prison warden, personnel, welfare and federal administrators have similar interests in the sense that on none of the 34 occupational interests do they differ significantly.[8] Forest service administrators do not differ from federal in any respect and differ very little from the other four. The first five, and possibly the six, may be viewed as the central core of public administrators as far as interests go. The remaining ten sub-groups may be classified according as they differ from them.

Six of the remaining ten sub-groups, i.e., engineer, chemist-physicist, statistician, city manager, public health and forest service supervisor, differ from the foregoing in having stronger scientific, production and mechanical interests (Groups I, II, III and IV) and weaker social service, public accounting, clerical and sales interests (Groups V, VII, VIII and IX). The other four sub-groups, i.e., taxation, comptroller-finance, city school superintendent and recreation, differ from the remainder in reverse manner from that of the six.

The writer believes that the relatively small differences between certain sub-groups is because they are heterogeneous groups. They are heterogeneous because society has not yet a clear conception of what is to be done in such positions and consequently permits a great variety of men to engage in it. As time goes on improvement along such lines will take place and it is likely that the sub-groups will become more homogeneous.

Our conclusions regarding federal administrators are misleading. Most of these men have an appreciable number of high interest scores but the averages of such scores are relatively low because of the heterogeneity of men engaged in many different activities. It is more or less futile to talk about federal administrators as such, even to talk about administrators in a single department, such as the Department of Agriculture, for they are engaged in too many different activities. Interests are related to activities. If the interests of men engaged in a wide variety of activities are averaged, low interest scores are inevitable.

Received January 25, 1946.

References

1. Strong, E. K., Jr. *Vocational interests of men and women.* Stanford University Press, 1943.
2. —— Interests of public administrators. *Pub. Pers. J.*, 1945, **6**, 166–173.
3. —— The interests of Forest Service Men. *Educ. & Psychol. Meas.*, 1945, **5**, 157–171.
4. —— Senior and Junior Public Administrator. *J. appl. Psychol.*, 1946, **30**, 55–71.
5. Tilton, J. W. The measurement of overlapping. *J. educ. Psychol.*, 1937, **28**, 656–662.

[8] There is one exception—welfare scores 8 higher than federal on minister interest giving a critical ratio of 3.3 and 75 per cent overlapping.

Reliability and Validity of the Kopas Wage Earner Battery of Tests *

Brent Baxter

Owens-Corning Fiberglas Corporation, Toledo, Ohio

Industrial users of aptitude tests are anxious to find tests which are especially easy to administer and score and that have a "practical" or face-validity appearance. Industry as a whole has not been well acquainted with the concepts of reliability and validity, and thus have not been greatly concerned with these aspects when selecting its tests. It is important to psychologists generally, therefore, to make these concepts better recognized and accepted in order that tests do not fall again into disrepute in industry as they tended to do after the testing boom that followed World War I. This article reports data obtained on a test which has recently been introduced to industry and accepted rather widely without adequate data on reliability and validity.

The Kopas Wage Earner Battery (Copyright 1943) has been developed by Joseph S. Kopas and marketed through the Personnel Test Company. The unique features of these tests are that the test questions appear on a large board (12″ × 16″) and the examinee's choices are indicated by his turning small pointers on a dial. The tests are scored by noting the position of the pointers on the back of the board where the correct dial positions (answers) are marked. This method is novel and appealing to the examinee and provides a very easy procedure for scoring. It also avoids the stigma of being called a "paper and pencil" test, a term which is sometimes used to imply in industry that the test is impractical for hourly employees.

While these advantages are offered for the Wage Earner Tests, few data are available concerning their reliability and validity. The manual written by Kopas describing the origin, development, and use of the tests presents brief sketchy data on validity and no information on reliability. Correspondence with companies reported as users of the tests revealed little conclusive data on validity and no data on reliability.

* This article is a "prior publication," the author paying complete costs. The scheduled 80 pages per issue is thereby increased by the corresponding amount; thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted and printed in their regular turn.

Reliability

To obtain some information in the latter area, a study was undertaken to obtain data on the split-half reliability of the six basic subtests of the Wage Earner Battery then in use at one of the plants of the Owens-Corning Fiberglas Corporation.

The major difficulty in obtaining split-half reliability data for these tests (and probably one reason why data were not already available) is that in the normal testing procedure, no record is made of the responses to individual items. In scoring the tests only the total number right as visually counted is marked on the test record sheet. All of the dials are then turned to a neutral position, destroying the record of item responses. To obtain split-half scores, it was necessary to establish special scoring procedures for each of the tests. The splitting of each test was made in consideration of the similarity of the test material in each half and the ease of counting the halves. Since the tests have no time limits, this did not force the use of the usual odd-even procedure which would have been difficult to score. The tests were scored as follows:

	First Half	Second Half
Test A—Mental Alertness	1–12	13–24
Test B—Background in Math. and Science	1–17	18–35
Test C—Personal Characteristics	1–6	7–12
	13–18	19–24
	25–30	31–36
	37–42	43–48
Test D—Interest in Routine Work	1–16	17–32
	33–48	49–64
Test E—Emotional Stability	1–12	13–24
	25–36	37–48
Test F—Ambition	1–24	25–48

Scores for the split halves were obtained for 100 applicants for hourly positions to whom the battery of tests was usually given. The test administrator noted the score for each half before turning the pointers to a neutral position for the next examinee. From these data a product moment correlation was calculated for each test. This was corrected for double length, the results being shown in Table 1. This table also shows the reliability coefficients as determined from the total score distributions by the Kuder-Richardson formula for Case IV.

None of the reliability coefficients are as high as might be desired for individual prediction. Tests C and F have the highest reliabilities and compare favorably in this respect to many other measures of personality.

The reliability of tests A, D, and E are so low that one's estimate of an individual's true score would be very inaccurate if made from these tests.

This pattern of low reliability springs from the author's concepts of the place of reliability in current testing. He writes,[1] "It is difficult to construct short tests that will be statistically reliable and valid. On the other hand, it is also difficult to validate a test that would be satisfactory to a psychologist while personnel work remains at its present level. It seemed logical, therefore, that it would be better to lean in the direction of the personnel supervisor's concept [2] of a good test so that he would accept and use tests and thereby improve his personnel program sufficiently to make available better criteria by which tests can be validated." It would appear from the data reported in Table 1 that in the case of the

Table 1

Coefficients of Reliability for Kopas Wage Earner Battery of Tests

Test	No. of Items	Split Half Reliability (Corrected)	Kuder-Richardson Reliability
A	24	.59	.69
B	35	.77	.76
C	48	.87	.86
D	64	.59	.58
E	64	.58	.69
F	48	.88	.82

majority of these tests, the "leaning" towards unreliability is so great that the personnel supervisor's program is not likely to be improved thereby. The validity of the true scores for the tests would have to be nearly perfect before the tests would be useful predictors. This is unlikely for the kinds of abilities claimed to be measured by the tests.

Validity

An interesting and novel slant on the use of tests is presented by the test author. "Because of the difficulties mentioned in developing aptitude tests" (need for criteria and the large number of separate jobs needing prediction) "and because of the need for additional information even when they are used, it was decided that a more practical approach

[1] *The development and use of tests in modern personnel programs* by Joseph S. Kopas (mimeographed), 1942.

[2] "Short, no timing, self-administering, self-scoring, self-rating, easily interpreted, and practical."

to the problem of testing factory workers would be exploring the testing method *as a means of securing information rather than as a way of checking people for aptitudes* required for specific jobs. The reasons for that conclusion are rather apparent. Seldom can an interviewer spend more than fifteen minutes per interview if he is to do the volume of work necessary in a typical personnel program. As a result whatever decisions have to be made regarding selection, placement, and training of new applicants are made on the basis of the information that is obtained by the short interview. Extending the interview is one approach to the problem, but it is costly and impractical. Supplementing the interview by less expensive methods of obtaining information would seem to be a much more economical method. The testing method, because it obtains information in an objective and organized manner when successful, should go a long way in getting the information necessary to do a good job of personnel work." [3]

The Wage Earner Tests were designed, therefore, to "give information about the applicant" rather than measure any specific abilities or aptitudes. In some cases, they have been sold as "interviewing aids" rather than as tests. By using this manner of speaking, however, it does not seem that the author is thereby excused from showing that the "information" is both accurate and relevant. By calling test results "information" it does not get around the need for criteria. While it may be extremely difficult to obtain a sound criterion, the advancement of scientific personnel programs depends upon this. To give the interviewer pertinent-looking information which in fact is irrelevant and inaccurate only builds a false sense of security in the soundness of the personnel action which he advises. The industrial use of tests will grow soundly only with adequate validation procedures.

The need for validation is especially important when claims are made for measuring new aspects of the individual. This is true of Tests C, D, and F of the Wage Earner Battery. The writer knows of no other tests in industrial use which purport to tap objectively the applicant's interest in routine work or his ambition and drive. While in many instances it would be highly desirable to have such information about the applicant, it is necessary first to be reasonably sure that the data thus collected are true. Too often test users, especially in industry, are willing to accept the test name as proof of what is measured without asking for evidence. For example, Test A is called Mental Alertness and is said to deal with the "ability of the individual to think in mechanical terms and to learn quickly and easily." The Test consists of 24 questions about eight pictures of gears and pulleys. This material is quite different from what

[3] *Ibid.*

is traditionally considered as content for mental alertness tests and no evidence is offered to show what it does measure. Thus interpretation in light of the test name and description is likely to be misleading.

Summary

1. The split-half corrected reliability coefficients obtained for the subtests of the Kopas Wage Earner Battery range from .58 to .88, half of them being less than .60.

2. The test author's concept of test validity is discussed and the need for the "long hard road" of careful validation studies is reemphasized.

Received November 5, 1946.

Luckiesh and Moss on Reading Illumination

Arthur C. Hoffman

Tufts College

In specifying optimal reading conditions, one of the crucial considerations is the level or intensity of illumination required for efficient reading. Luckiesh and Mass (**9**, p. 350) recommend:

> *50 to 100 foot-candles* for severe and prolonged tasks such as proofreading, difficult reading, etc.
> *20 to 50 foot-candles* for moderately critical and prolonged tasks, such as clerical work, ordinary reading, etc.
> *10 to 20 foot-candles* for moderate and prolonged tasks of office and factory, and when not prolonged, ordinary reading, etc.

However, considering other discussions of their results which Luckiesh and Moss have published, the above would seem to be their more conservative statements. The tone of their writing would suggest that they prefer the following recommendations:

> It has been definitely established that intensities of illumination as high as 100 foot-candles are desirable for such a visual task as reading ordinary black print upon white paper (**8**, p. 273).
>
> Studies have usually been limited to a range of illuminations up to 100 foot-candles, but since white backgrounds have been used, it appears that much higher illuminations would be advantageous where backgrounds have a lower coefficient of reflection (**13**, pp. 6–7).
>
> From the foregoing brief analyses it is indicated that the ideal level of illumination for reading even large type excellently printed is above 100 foot-candles. Therefore, ten foot-candles on 8-point type is a very conservative standard of visibility (**9**, p. 343).

As has been pointed out in the experimental literature, these recommendations are somewhat at variance with the conclusions drawn by practically every other investigator of this illumination problem (e.g., **5**, **18**, **19**). It might be well, then, to review the justification for their recommendations offered by Luckiesh and Moss.

In the opinion of the present author, their technique of analysis and their interpretation of their experimental data are questionable. It is the primary purpose of this paper to reinterpret their results and, where possible, to present a statistical reanalysis of the experimental data published by them. Where the raw scores for each subject have been available, the present author has reanalyzed their data for reasons to be given below. Because of the smallness of the samples, Fisher's *t*-test technique has been applied with a 5 per cent level of confidence.

Percentage Technique

When the results of the reading-illumination experiments performed by Luckiesh and Moss are reviewed, it is noted that in the majority of instances these authors use a type of percentage score technique to evaluate or express their findings (see Table 1). Their general method is to express a particular raw score (usually derived from an experimental situation regarded as standard or control) as 100 and then to equate all the other raw scores to 100 in the same ratio as they are related to the particular criterion score. For example, if the mean scores from three experimental situations to be compared were 20, 21, 25; by this method the 20 would be expressed as 100, and the other two scores as 105 and 125. Comparisons, and the conclusions drawn from them, appear to have been made with reference to the differences noted in these percentage scores and not with reference to the original raw scores.

Table 1

Average Number of Blinks Occurring During the First and Last Five-Minute Periods of an Hour of Reading Under Levels of Illumination of 1, 10, and 100 Foot-candles

(After Luckiesh and Moss, 11 and 12)

	1 ft-c		10 ft-c		100 ft-c	
	first	last	first	last	first	last
Number of blinks	35	60	35	46	36	39
Relative rates of blinking	100	171	100	131	100	108
Per Cent Increase	71.5 ± 5		31.4 ± 3		8.3 ± 1	

Sufficient raw data from the experiments of Luckiesh and Moss were not always available to the present author to enable him to present a comparison of the conclusions to be drawn from analysis of raw data with the conclusions drawn by Luckiesh and Moss from their per cent scores. It should be noted, however, that this percentage technique is open to serious question for reasons in addition to the notorious unreliability of percentage scores.

As may be shown by the examples below, this technique has the function of magnifying differences if the original raw scores are below 100 in absolute magnitude, or of minimizing differences if the absolute magnitude of the original scores is above 100. And this is true, even though the absolute magnitude of the differences between raw scores is the same in both cases. For example, if two means of raw scores are 50 and 60, and two others are 150 and 160 (both sets showing a difference of 10), this percentage technique would express them as 100 and 120, and 100

and 106.7. Now a difference of 20 may be noted in the first instance and one of only 6.7 in the second.

Similarly, when scores of different ocular functions are compared by this technique (e.g., rate of blinking under 1 and under 10 foot-candles of illumination and rate of reading under the same illuminations) as Luckiesh and Moss have done, differences between raw data in the case of both measures may be the same, but the percentage expression of these differences varies depending on the relation of each raw basic or criterion value to 100. For example, if the original means for one measure are 10 and 20, and the means for the second measure are 40 and 50 (both showing an apparent difference of 10), the percentage expression of these scores would be 100 and 200, and 100 and 125. Basing a conclusion solely on the apparently greater difference in the percentage measures of the first instance overlooks the fact that the observed differences may be just an effect of the derivation and not necessarily a demonstration of statistical significance.

The most serious objection to the technique of Luckiesh and Moss is that the percentage values so derived have no easily determined standard error. They, therefore, lack a valid criterion with which to interpret or evaluate the significance of the differences observed, i.e., to determine whether the differences are due to some other factors than just chance fluctuation. It has been shown (6) by the present author that when increments or decrements in several measures of reading performance are expressed in terms of per cent, several such derived scores can all indicate significant change in performance even when they differ in absolute magnitude; also, several per cent scores of the same magnitude may not all be indicative of significant change. There is no definite clue to these facts in the magnitude of the per cent scores themselves. The significance of differences must, therefore, be determined by further, appropriate, statistical techniques. The conclusions to be drawn in answer to the present problem of illumination should, of course, be based as much as possible upon this determinable significance of observed experimental differences—hence, the reanalysis by the present author of raw data published by Luckiesh and Moss.

Indicated Nervous Muscular Tension

Luckiesh and Moss suggest that the output or decrement overtly manifested in an experiment may not be a complete or an adequate criterion of the effort expended in useful work or of the strain of the work on the subject. In discussing their experiment on indicated muscle tension they state:

"As the intensity of illumination is increased, or as seeing is improved by other means, the production of useful work (and the energy expended in useful work)

increases and eventually reaches an optimum. . . . It is a matter of everyday experience that nervous tension and fatigue commonly cause a tightening of certain muscles. . . . It may also be observed that muscular energy is often wasted as a result of strain or fatigue which originates remotely from the place where it manifests itself. . . . In this case, an inadequacy in conditions for seeing is manifested in the useless expenditure of muscular energy. . . . The muscular strain, using the term in its usual sense, produced by fatigue or strain is referred to as *nervous muscular tension* in the present discussion. Measurements of this factor are designated as *indicated nervous muscular tension*, since we have no means of knowing the completeness or tne accuracy with which this criterion appraises or measures these internal psycho-physiological reactions.

"The method evolved and used in the research consisted in measuring the changes in muscular tension as registered by the fingers of the left hand which rested naturally and lightly upon the knob of a concealed key" (**7**, pp. 541–543).

The average data (geometric means, whose application to this situation could be questioned) show that the indicated nervous muscular tension, developed incidental to reading, decreased from 63.2 grams, to 54.1 grams, to 43.0 grams as the intensity of illumination was increased from 1 to 10 to 100 foot-candles. Analysis by the present author of the arithmetic mean differences in indicated tension under these conditions has shown that these differences may be regarded as highly significant. From their data Luckiesh and Moss conclude:

"It is apparent that the differences between psycho-physiological reactions resulting from reading under different levels of illumination are of such definiteness and magnitude that they are capable of creating muscular strains which are measurable. . . . This *indication* of the waste of human energy due to inadequate levels of illumination may be considered as a conservative appraisal of the actual conditions. Certainly, it is measured at a point remote from its origin. Apparently it can indicate nothing else than an effect of greater ease in performing a relatively easy visual task under intensities greatly above the threshold requirements" (**7**, p. 553).

Overlooking the knowledge that muscle tension may in some instances facilitate behavior (**2**, **3**), the statement of Luckiesh and Moss might be taken at face value that muscle tension is an indication of wasted energy. The argument would appear to be that this tension should be reduced or entirely eliminated if possible. The experiment just reviewed demonstrates that muscle tension (whatever its causes or functions turn out to be) decreases as illumination is increased. It is to be noted, however, that data are offered (**7**) which show a significant decrease under 10 foot-candles, *and also* a significant decrease under 100 foot-candles. It would seem to the present author that the only conclusion that can be drawn from this evidence is either that tests of muscle tension should be made while reading under still higher and higher intensities until the tension becomes minimal or shows no further significant decrease; or, for the present, to recommend some illumination considerably above 100 foot-

candles. Luckiesh and Moss would seem to lean toward the latter alternative, judging from their published recommendations.

It may be interesting, however, to consider the recommendations of Luckiesh and Moss in relation to certain other findings. It is reported (e.g., **1, 14**) that glare interfers with visual acuity. It is known that the possibility of glare interference begins at around 25 foot-candles (**19**). Ferree and Rand (**4**) report that glare may interfere at even lower intensities when artificial light (e.g., Mazda lamp), rather than daylight, is used for illumination due to the yellow color and the relatively poor diffusion of artificial light. At 100 foot-candles or more the possibilities for glare and therefore an adverse affect on visual acuity would hardly be remote. It is also known (**16**) that pupillary diameter or size depends in large measure on the intensity of illumination incident upon the eye; the pupil constricts in the presence of bright light, especially if that light is close to or directed into the eye as would be the case while reading. The brighter the light, therefore, the smaller the pupil. At extremes of constriction, however (as well as at extremes of dilation), visual acuity is impaired. The above would suggest that the intensity of illumination is not the only factor to be considered in prescribing for easy reading, as Luckiesh and Moss themselves point out.

The results of the experiment just considered are, thus, inconclusive. And the question being raised is whether or not too high a critical intensity has not already been reached at the levels recommended by Luckiesh and Moss or concluded by implication from this muscle tension experiment.

The Reading Rate Criterion

The speed of reading, usually expressed as rate of reading, has frequently been used as the criterion variable in ease-of-reading experiments. Luckiesh and Moss object to many of these experiments (and hence the conclusions drawn from them) on the grounds that rate of reading is a poor criterion with which to test visibility or ease of reading. For example,

> "The results of a series of carefully controlled investigations, involving diverse visual variables, reveal that the normal rate of reading is an insensitive indicator of readability as compared with rate of blinking. Furthermore, the appraisals of readability by the two criteria are frequently in disagreement. Whether or not the criterion of involuntary blinking adequately appraises all factors involved in readability, it appears to be far more significant, from both theoretical and experimental viewpoints, than the criterion of rate of reading" (**12**, p. 269).

It is difficult for the present author to evaluate the above statement with data as presented by Luckiesh and Moss. It is a question of whether this disagreement is a valid function of the carefully controlled investiga-

tions (as it might be) or of the technique of presentation and evaluation of data.

Subject data on *rate of reading* have been published by Luckiesh and Moss (9) from one experiment in which fourteen subjects read for thirty minutes in each case under 1, 10, and 100 foot-candles of illumination. From another experiment, subject data are presented on the *rate of blinking* while reading for five minutes under each of the same illumination intensities. The mean scores for rate of reading were, respective to illumination intensity, .716, .728, and .747; for rate of blinking the corresponding mean scores were 34.6, 34.6, and 35.4. Recalling the percentage technique previously described, these data would be reported as 100, 101.7, and 104.3 for rate of reading; and 100, 100, and 102.3 for rate of blinking.

Analysis by the present author of the differences between means permits the conclusion that reading rate under 1 foot-candle was not significantly different from that under 10 foot-candles, nor was that under 10 foot-candles significantly different from reading rate under 100 foot-candles. But rate of blinking was also not significantly different under 1 foot-candle as opposed to 10 foot-candles, under 10 foot-candles as opposed to 100 foot-candles, nor under 1 foot-candle as opposed to 100 foot-candles. It will be noted that the results of these particular experiments hardly support the recommendation of 100 foot-candles as opposed to lower intensities. But more important at the moment is the fact that conclusions to be drawn from the rate of reading results are the same as those to be drawn from the rate of blinking—in short, no distinction in the effectiveness of the two criteria.

The above consideration would suggest that the contention of Luckiesh and Moss is at least an overgeneralization. The preceding statement, of course, does not necessarily imply that the blinking rate cannot be used as a criterion with which to investigate the illumination problem at hand. It is also recognized that other research by Luckiesh and Moss to support the contention that the rate of reading is an insensitive indicator as compared with rate of blinking might warrant such a distinction if it could be properly analyzed for significance of differences in the case of each criterion.

The Rate of Blinking Experiment [1]

Luckiesh and Moss report an experiment (11) which they may regard as crucial, judging from the frequent use they have made of the data.

[1] Since writing this critique, two research studies throwing doubt on the validity of rate of blinking have been published by M. A. Tinker: Reliability of blinking frequency employed as a measure of readability. *J. exper. Psychol.*, 1945, **35**, 418–424, and Validity of frequency of blinking as a criterion of readability. *J. exper. Psychol.*, 1946, **36**, 453–460.

Eleven subjects read for an hour at a time under 1, 10, and 100 foot-candles respectively. The blinking rate was determined for the first and the last five minutes of each such hour of reading. Luckiesh and Moss accept the tentative hypothesis of Ponder and Kennedy (**15**) and others that the rate of blinking is affected by or may be indicative of the strain or tension caused in an individual by, among other things, the difficulty of a task. The assumptions, then, upon which this experiment was based seem to be that illumination intensity while reading is inversely related to tension (the lower the illumination, the greater the tension) and that the degree of this tension will be indicated in some measure by the rate of blinking. The experiment, of course, was not designed to test these assumptions; but accepting them, it had as its purpose the determination of the illumination intensity requisite for facile reading.

The mean data and the "per cent increase" reported by Luckiesh and Moss are shown in Table 1. Noting the decrease in the "per cent increase" in blinking (71.5, to 31.4, to 8.3) as illumination was increased,

Table 2

Analysis of the Differences Among the Averages Shown in Table 1

	Five-minute periods compared	*t*-ratios
Group I.	a. 1 ft-c; first vs. last	2.654 *
	b. 10 ft-c; first vs. last	2.450 *
	c. 100 ft-c; first vs. last	2.410 *
Group II.	a. 1 ft-c first vs. 10 ft-c first	0.000
	b. 10 ft-c first vs. 100 ft-c first	0.830
Group III.	a. 1 ft-c last vs. 10 ft-c last	2.594 *
	b. 10 ft-c last vs. 100 ft-c last	2.251 *

* Significant at the 5% level of confidence.

they conclude in favor of 100 foot-candles. The significance of the differences between the mean blinking rates shown in this table has been determined by the present author using the data presented on page 230 of *Science of Seeing* (**9**). The *t*-ratios evaluating each difference are shown in Table 2. They are listed according to the comparisons made between first or last five-minute intervals under 1, 10, or 100 foot-candles of illumination. The three ratios labelled Group I test the differences between the mean blinking rates of the first and the last five minutes of reading under each of the three illumination intensities. The Group II ratios compare the first five-minute intervals under each intensity. The last five-minute intervals of each level of illumination are compared by the Group III ratios.

The ratios in Group I of Table 2 compare the rate of blinking at the

beginning and end of an hour's reading under each of the three illumination intensities. It will be noted that in all three instances the hypothesis is tenable that blinking significantly increased. At this point the same comment can be made of these results as was made of the experiment on indicated muscular tension previously reviewed. Since the criterion variable was found to change significantly in all three circumstances, and in a direction regarded as disadvantageous, no crucial or optimal intensity is necessarily established by this experiment.

It may be difficult to justify the design of this experiment in keeping with its purpose. That purpose was to test an effect of the illumination variable. During the *first* five minutes of reading under each illumination intensity, it is reasonable to suppose that the most salient factor varied was the illumination intensity. These several sample periods are compared by the Group II ratios, neither of which may be regarded as significant. Thus no significant change in the rate of blinking from one illumination level to the next is indicated. The intensity of illumination, then, does not seem to be a significant variable as tested in this manner.

In their interpretation, however, Luckiesh and Moss would seem to consider the last interval of the hour's reading as well as the first. For this reason the Group III ratios are presented. When the blinking data after an hour's reading under 1 foot-candle are compared with the blinking rate after reading under 10 foot-candles, the latter are found to be significantly lower (Group III-a). But when the last five minutes of reading under 10 foot-candles is compared with the last five miuntes of reading under 100 foot-candles (Group III-b), a significant difference in blinking rate is also found. An experiment by the present author (6) has shown that there are significant changes in the rate of blinking and in other eye-movement measures of reading under *constant* illumination, *which changes already occur within an hour of continuous reading.* This, the Group I ratios, and the above led the present author to wonder whether or not Luckiesh and Moss have demonstrated a "work decrement" of some sort.[2] It may be that what has been observed here is the effect of a work factor coupled with an illumination variable.

This consideration focusses special attention on blinking performance during the last 100 foot-candles five minutes as compared with the *last 10 foot-candles five minutes* (Group III-b) and with the *first 100 foot-candles five minutes* (Group I-c). The even slight difference in magnitude of the two *t*-ratios evaluating these comparisons may suggest that performance at the end of an hour's reading under 100 foot-candles is more

[2] The term decrement is not used to mean that the number of blinks decreased, for it will be noted that they actually increased. Rather the term is used to imply a change of a presumably deleterious nature—the higher the rate of blinking, the worse the reading situation is assumed to be.

like that at the end of reading under 10 foot-candles than it is like that at the beginning of reading under 100 foot-candles. This implies that decrement or the work-in-time factor was more important than the illumination variable in influencing the data.

An additional comment should be made, however. The difference in magnitude of the two *t*-ratios in Group III may suggest that the decrement was less while reading under 100 foot-candles than while reading under the lower intensities of illumination. The difference between the last five-minute periods of 100 foot-candles and 10 foot-candles (Group III-b) is seemingly less than the difference between 10 foot-candles and 1 foot-candle (Group III-a). All the ratios in Group I may be regarded as significant; if also the second ratio (b) in Group III had been insignificant, it could then be concluded that decrement under 100 foot-candles is no different from decrement while reading under 10 foot-candles. Therefore, 10 foot-candles would be just as good as 100 foot-candles as far as minimizing 'fatigue' is concerned. But, this comparison of the last five minutes under 10 foot-candles with the same period under 100 foot-candles (Group III-b) may be regarded as significant. Even though the question of whether some intensity less than 100 and more than 10 foot-candles would be as satisfactory is still left open, the above consideration permits the conclusion that reading is easier, or rather is less difficult, under 100 foot-candles than under 10 foot-candles. This finding, then, is the supporting nugget that has been sought in this discussion since the recommendations of Luckiesh and Moss were first states.

It may be of interest to statistically minded readers, however, to know that none of the *t*-ratios reported in Table 2 in connection with the above blinking experiment would be regarded as significant if a 1 per cent level of significance had been chosen. It seems to the present author that such a level of confidence could have been quite legitimately chosen in view of the contrary findings reported by other experimenters. Before accepting the recommendations of Luckiesh and Moss, which are at variance with the majority of the other conclusions drawn, one would wish to be very confident that the recommendations were justified. Ratios non-significant at the 1 per cent level of confidence would, of course, mean that no significant differences had been demonstrated by Luckiesh and Moss. The task of interpretation would be very difficult with the evidence at hand, since the above findings would allow many conclusions. Few, if any, however, would support the contentions of Luckiesh and Moss.

In reviewing his own and related research Tinker has stated:

"The experiments of Luckiesh and Moss have led them to recommend what seem to be excessively higher light intensities for visual tasks. . . . Although

all indications suggest that the critical level for reading 10- to 12-point type is around three to four foot-candles (most certainly this critical level is below ten foot-candles) it is suggested that any specifications made should provide a margin of safety. The conclusion is that ten to fifteen foot-candles should provide hygienic conditions when one's eyes are normal and print is legible" (17, p. 570).

With this conclusion the present author would concur. In fact, if he were willing to attempt a best generalization from the experimental evidence available, he would recommend about 15 foot-candles for the ordinary studying and reading situation.

Received March 17, 1946.

References

1. Beacher, L. L. Possibilities of glare reducing visual acuity. *Amer. J. Optom.*, 1932, **9**, 338–340.
2. Bills, A. G. Facilitation and inhibition in mental work. *Psychol. Bull.*, 1937, **34**, 286–309.
3. Courts, F. A. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, **39**, 347–367.
4. Ferree, C. E., & Rand, G. Work and its illumination. *Person. J.*, 1940, **19**, 93–98.
5. ———. Optimum working conditions for the eye. *Sight-Sav. Rev.*, 1940, **10**, 3–12.
6. Hoffman, A. C. Eye-movements during prolonged reading. *J. exper. Psychol.*, 1946, **36**, 95–118.
7. Luckiesh, M., & Moss, F. K. A correlation between illumination intensity and nervous muscular tension resulting from visual effort. *J. exper. Psychol.*, 1933. **16**, 540–555.
8. ———. Relative visibility of print in terms of illumination intensity. *Sight-Sav. Rev.*, 1935, **5**, 272–280.
9. ———. *The science of seeing.* New York: D. Van Nostrand Company, 1937. Pp. 548.
10. ———. Frequency of blinking as a clinical criterion of ease of seeing. *Amer. J. Ophthal.*, 1939, **22**, 616.
11. ———. A summary of researches involving blink-rate as a criterion of ease of seeing. *Illumin. Eng.*, 1940, **35**, 19–32.
12. ———. Criteria of readability. *J. exper. Psychol.*, 1940, **27**, 256–270.
13. ———. Illumination and visual efficiency. *Amer. J. Optom.*, 1942, **19**, 5–15.
14. Melian, A. La fatiga ocular: sus factores y caracteristicas. (Ocular fatigue: its factors and characteristics.) *Psicotechnia*, 1940, **1**, 81–100.
15. Ponder, E., & Kennedy, W. P. On the act of blinking. *Quart. J. exper. Physiol.*, 1927–28, **18**, 89–110.
16. Reeves, P. The response of the average pupil to various intensities of light. *J. Opt. Soc. Amer.*, 1920, **4**, 35–44.
17. Tinker, M. A. The effect of illumination intensities upon speed of perception and upon fatigue in reading. *J. educ. Psychol.*, 1939, **30**, 561–571.
18. ——. Illumination intensities preferred for reading with direct lighting. *Amer. J. Optom.*, 1944, **21**, 213–219.
19. Troland, L. T. An analysis of the literature concerning the dependency of visual functions on illumination intensity. *Trans. Illumin. Eng. Soc.*, 1931, **26**, 107–196.

Some Evidence on the Validity of the Cardall Test of Practical Judgment

Edwina Hogadone and Leo F. Smith

Rochester Institute of Technology, Rochester, New York

Laymen and psychologists alike have long felt the need for an instrument which would measure common sense or practical judgment. One instrument which has been constructed in the effort to satisfy this need is the Test of Practical Judgment by Alfred J. Cardall. The purpose of this test as stated on page 1 in the preliminary manual is, "to measure the element of practical judgment as it operates in everyday business and social situations. This measurement is unique and statistically independent of such factors as intelligence, academic and social situations. Practical judgment is an important consideration in the selection of executives, supervisors, salesmen, those who are working in contact occupations, and those responsible for planning and supervising work. Schools use the test as a measure of 'social maturation.'" [1]

In order to determine what practical judgment consisted of Cardall asked a number of personnel men and psychologists what they understood by practical judgment and the consensus was (p. 1), "That practical judgment consists of the recognition of possible alternatives of action and the ability to select the best. It may be observed that in ordinary situations a number of alternatives spring instantly to mind, the element of judgment operating when one course of action is chosen over others. Upon this rationale the present instrument was constructed." In the manual the following statement is made about the reliability (page 2), "The reliability of this test, computed on a homogeneous group, is .88 by the method of split-halves and .86 by the Kuder and Richardson *r* formula. In other samplings, but on a smaller number of cases in which the range of ability has been greater, the reliability coefficient has been materially higher." In neither of these samples does Cardall indicate the number of individuals involved.

As evidence of the validity of the instrument the author makes the following statement (p. 3), "In one study recently completed in a public utility company, test results correlated .56 with 'value to the company, in the accounting department as determined by pooled supervisors'

[1] Cardall, Alfred J. *Preliminary manual for the test of practical judgment.* Chicago: Science Research Associates. 1942.

rating." Here again the number of individuals involved was not indicated.

Purpose of the Study

It is the purpose of this paper to report a brief study of the Test of Practical Judgment. For many years the Educational Research Office has worked closely with the Retailing Department in the effort to improve techniques for predicting success in this course. It was in the hope that the TPJ might be of some value in this regard that the study was made. During the spring of 1945 the test was administered to all of the seniors (N = 33) in the Retailing Department. During the spring of 1946 the TPJ was administered to the present junior and senior classes (N = 16 and N = 42 respectively). All of these 91 students were enrolled in a three-year cooperative course during which they had alternated at four-week intervals between academic work at the Institute and employment in the department stores of western New York. Prior to each administration of the test the supervisor of the department had utilized the foregoing definition and rated all of the students on their practical judgment as exhibited in school and cooperative job situations. These ratings were made on a five point scale, namely, superior, good, average, poor and unsatisfactory. The test was then administered, scored and correlations run between the supervisor's rating and the test scores. In addition a correlation was run between the supervisor's ratings and the gross scores of the 1941 Edition of the American Council on Education Psychological Examination, College Form. The ACE had been administered to all of the students prior to their entrance to the Institute.

The authors are well aware of the unreliability of ratings and the difficulty of estimating a student's "common sense" or "practical judgment." Because of this administrative set-up of the Institute, however, the relation between the supervisor and the student is extremely close. The supervisors are responsible for counseling students in their departments and have at their command a "behavior journal" [2] which brings together the results of classification examinations, incidents of student behavior noted by faculty members, reports from cooperating employers, summaries of conferences held with students, and periodic syntheses of individual adjustments and personal needs as revealed by the continuous picture obtained from the journal. In view of the foregoing, it appears probable that the supervisor's ratings are more reliable than those normally obtained from faculty members or foremen in industry.

[2] Jarvie, L. L., and Ellingson, Mark. *A handbook on the anecdotal behavior journal.* Chicago: University of Chicago Press, 1940.

Results

The correlation between the raw scores on TPJ and the supervisor's rating of "practical judgment" for the 91 students in this study was 0.24. According to Fisher[3] this is significant between the 5% and 2% level and corresponds to a z[4] of 0.2403.

The correlation between the raw scores on the ACE and the supervisor's rating of practical judgment was 0.28. This is significant at the 1% level and corresponds to a z of 0.2806.

In an effort to determine if there was a significant difference between the correlations thus obtained Fisher's test for the significance of the difference between two observed correlations[5] was applied. In this case t proved to be .267. According to Fisher (Table IV, p. 177) P is between .7 and .8 and we cannot assume with any degree of assurance that the difference between the two correlations is not due to chance. Apparently for the group of students included in this study the TPJ is not quite as good a measure of practical judgment as rated by the supervisor as is the ACE. The difference, however, is not significant.

Interestingly enough, the correlation between the raw scores on the TPJ and the gross scores of ACE was 0.22. Although the Cardall test is not a good measure of practical judgment in this particular situation, neither is it highly correlated with academic aptitude as measured by ACE.

Summary

1. This study compared the results of raw scores made by 91 students on the Cardall Test of Practical Judgment and gross scores made on the ACE with supervisor's rating of the students' practical judgment as exhibited in class, cooperative job and extracurricular activities.

2. Results indicate that the ACE correlated slightly higher (0.28) with the supervisor's rating than did the TPJ (0.24). According to Fisher's z transformation for determining the significance of the difference between correlations this difference is not significant.

3. Although the TPJ is not a good measure of practical judgment in this particular school situation, neither does it correlate highly with academic aptitude (0.22) as measured by the ACE.

Received January 7, 1946.

[3] Fisher, R. A. *Statistical methods for research workers.* 7th Ed., 1938, Table V.A., p. 214.

[4] *Ibid.* Table V.B., p. 215.

[5] *Ibid.*, p. 208–209 (Peters and VanVoorhis in *Statistical procedures and their mathematical bases*, 1940, p. 188, have presented the formula for determining the significance of the difference between z's in a way which is much easier to utilize than has Fisher).

Interests and General Educational Development

Arden Frandsen

Utah State Agricultural College

The assumption that motivation plays an important role in learning is both generally accepted by psychologists and educators and well supported by observations and experiments (**11**, ch. 16). In the review cited, Stroud writes: "Not only do motives arouse and direct behavior, create learning situations and define the solution, they are also intimately connected with the act or process of learning." And "Whether or not a pupil gets up his lesson depends upon the strength of the motive-incentive conditions appropriate thereto and upon the strength of competing motives and incentives." The quantitative degree of this relation between motivation and learning achievement is indicated in a study reported by Frandsen and Darke (**4**, p. 129). They found a correlation between high school teachers' and counselors' ratings on total motivation[1] and first year grade-point-ratios of 353 University of Utah freshmen to be .47. When intelligence, which correlated to the extent of .30 with the ratings on motivation, was partialled out, the correlation between motivation and achievement was still .41. Between similar ratings on motivation and the school marks of 65 high school students Turney has reported even higher correlations (**14**).

Among the motives which teachers rely upon to arouse, sustain, and direct learning activity (adapting to our purpose Shaffer's definition of motivation **9**, p. 84), appeal to interests is generally regarded as very effective. Pressey and Robinson, for example, observe that "To have an interest is to feel a hunger, an urge, a craving, a restlessness. And when one is in the grip of an interest, he has excitement, thrill, satisfaction, and a sense of expanding well-being—or letdown, discouragement, frustration—according to whether the interest is satisfied or not" (**7**, p. 148).

As a consequence of this identification of interests with important learning motives and because of the well-established relation between motivation in general and learning achievement, it is perhaps only logical to expect that between vocational interest inventory scores and learning achievement in corresponding areas there would be found relatively close

[1] The measure of "total motivation" used in the study was ratings on the trait of "industriousness" from the American Council on Education Rating Scale.

relationships. If the degree of association between "total motivation" and achievement mentioned above were adopted tentatively as a standard, correlations between .40 and .50 might be anticipated. Most of the data reported to the present on this relationship, however, are quite contrary to this expectation. In a recent partial review (**5**) of this literature, which included the findings of seven different investigators, correlations between interest inventory scores and achievement in corresponding subject-matter areas of from .04 to .23 were reported, except in three instances of higher relationships in science and mathematics which will be referred to later. As typical of most of the results on this interest-achievement relationship, the specific data from a recent study by Bolanovich and Goodman are presented.

One month after beginning training in Purdue University in a program of mathematics, shop, drawing, A-c and D-c, radio manufacturing, communications, measurements, and electronics, 66 women ranging in age from 18 to 28, who had been selected by RCA for training, were given the Kuder Preference Record interest inventory. The correlations between total grades in the ten months course and the nine parts of the interest inventory were: mechanical, .09; computational, .18; scientific, .10; persuasion, .15; artistic, .08; literary, .02; musical, −.02; social service, −.03; and clerical, −14 (**3**, p. 320). The correlations between the mechanical, computational, and scientific parts of the interest inventory and grades in this mechanical-mathematical-scientific curriculum are far too low to confirm the hypothesis that variations in these interests (or motives) affected significantly the differences in achievement among the students in this group.

To rationalize such low interest-achievement relations with the assumption that interests as motives are intimately connected with learning, the writer of the review referred to suggested that "besides interest, several other motives—mastery in general, social approval, acsendancy in competition for good grades, security, etc.—are appealed to in all courses. And school marks are not sufficiently sensitive to detect the effects of the interest motive on achievement when all these other motives are operating simultaneously with interest. In a situation, however, where no other (extrinsic) motives were affecting study efforts, the presence or absence of the (intrinsic) interest motive should make an appreciable difference in achievement" (**5**, p. 7). It is assumed, of course, that interests, like such motives as mastery and social approval, are learned. Interest provoking situations have probably become conditioned stimuli which arouse, sustain, and direct activity as a consequence of their previous association with responses reinforced by satisfying activity, mastery, and approval. But, according to Allport (**1**, ch. 7)

such learned motives may become functionally autonomous. And it is probable, therefore, that abiding interests may serve to integrate learning experiences over long periods into effective patterns of achievement.

Professor E. K. Strong, prior to the statement quoted above, offered a similar and more explicit hypothesis. "Interests," he writes, "may not correlate to any great degree with achievement over a short period of time and yet may correlate significantly when achievement involves performance over a considerable period of time" (**10**, p. 18). In elaboration of this view, he observes that "When a student discovers he has mistakenly elected a course in which he has little interest, he will finish it about as well as other courses but he will not elect further courses of a similar nature." Thus in a single course where some students may be motivated by intrinsic interest, and others, lacking intrinsic interest, may be motivated equally intensely by extrinsic motives, the effect of interest on achievement would be masked by the effects of these other motives, at least as far as correlation analysis can reveal. To demonstrate the possible effects of interest on achievement, therefore, will require a situation in which the inferred masking effects of extrinsic motives are minimized. It is the purpose of this paper to present data on the relation of interests to learning achievement in just such a situation.

Problem

Students at the Utah State Agricultural College are required to complete before graduation 8 to 12 quarter hours in each of the following areas: biological science, physical science, social science, and language and arts, plus 5 hours of sophomore composition. They are advised to complete these "group requirements" in the freshman and sophomore years, but elections of courses among these areas are made freely. For the 81 sophomores for whom data are reported in this study, the amount of credit earned in the natural sciences (combination of biological and physical sciences) varied from none to 64 quarter hours and in the social studies it ranged from none to 44 quarter hours, depending upon the interests of the students.

This situation affords an opportunity to test Strong's hypothesis that interests "may correlate significantly (with achievement) when achievement involves performance over a considerable period of time." This inference follows as a result of the assumption, implicit in the hypothesis, that the amount of knowledge accumulated in an area will vary with the number of courses taken in it and further that the number of courses elected will depend upon the degree of interest. To measure such cumulative, long range achievement, the averages of marks in the separate college courses would not be suitable, because, besides the college courses,

high school courses, and many out-of-school activities would also contribute to a student's total knowledge in a field. It will require a test measuring the attainment of broad, general objectives for the whole program of general education in these areas.

The recently provided USAFI Tests of General Educational Development would seem well suited to the purpose. The authors of these tests say, "This type of test can thus require that an integrated body of knowledge be brought to bear on particular problems, without placing any undue premium upon the peculiar form or organization in which the student's ideas have been acquired or without penalizing him unduly for inability to supply any particular fact or set of facts where another will serve the same general purpose" (2, p. 5). In other parts of the manual the authors state that "The tests are accordingly designed to measure the extent to which *all* of the past educational experiences of the individual . . . have contributed to his general educational development." For some students, interest may be the dominant motive in acquiring and integrating the knowledge required for effective performance on these general tests. It is also recognized that scores on the tests would be affected to some extent by the extrinsic motives which were appealed to in all the separate courses taken by the students in the areas covered by the tests. But intrinsic interest rather than extrinsic motives probably determined to a large extent the number of courses taken in each area by the students as well as the amount of knowledge acquired through voluntary reading, attendance of lectures, museum experiences, and out-of-class discussions. Thus the extent to which interests may affect achievement should be masked much less by the effect of extrinsic motives on the learning measured by these tests than on the tests usually applied to measure achievement in a single course.

Data and Interpretations

The data include three measures: interests, relative number of courses taken in natural science and in social studies, and scores on tests of general, long-range achievement—all obtained on 81 sophomores at the conclusion of their sixth quarter at the Utah State Agricultural College. The interest test was the Kuder Preference Record. The number of courses in the fields of natural sciences and of social studies taken by each student in both high school and college were tabulated. To equate for some inequality in the total number of credits earned by different students, the percentages of credit earned in each of these two broad fields were computed. Since in the field of language the test used was restricted to measuring composition ability and also because prescription of composition courses in both high school and college limits student

electives in this field, the credits earned in the language and arts group were not tabulated. The measures of general, long-range achievement were the USAFI Tests of General Educational Deveopment in (1) Natural Sciences, (2) Social Studies, and (3) in Correctness and Effectiveness of Expression.

The intercorrelations among these three achievement tests, for our subjects, indicate that the tests are measuring some abilities common to all three tests, but also that some different functions are measured by each test. Scores on the Natural Science Test correlate with Social Studies and Correctness and Effectiveness .57 and .53, respectively. Between the latter two the correlation is .50.

1. *Relation between interests and long-range, general achievement.* To show the relation between interests and long-range achievement in corresponding interest-achievement areas, the correlations between the scores on these factors are presented in Table 1.

Table 1

Correlations between Kuder Interests and USAFI (College Level) Tests of General Educational Development for 81 Sophomores, 71 Women, 10 Men

Achievement Tests	Interests								
	Mechanical	Computation	Scientific	Persuasive	Artistic	Literary	Musical	Social Service	Clerical
Natural Sciences	.14	.31	.50	−.33	−.02	.17	−.10	−.30	−.16
Social Studies	.02	.05	.34	−.12	−.14	.31	−.09	−.37	−.21
C & E of Expression	−.03	−.02	.15	−.19	.06	.16	.15	−.37	−.51

Contrary to the zero or low correlations usually found between interest and achievement in *single courses*, some significant and logically consistent correlations between interest and long-range achievement may be observed in Table 1. The correlation of .50 between science interest and achievement in natural science is especially important. An examination of the scatter-diagram shows further that every person above average in science interest achieved a relatively high score on the general achievement test in natural sciences. The correlation would have been even higher than .50, except for the fact that several students with relatively low science interest scores also earned high science achievement scores. This suggests the masking effect of extrinsic motives referred to in the introduction, and which in this instance probably compensated for the lack of intrinsic interest otherwise required to sustain the effort necessary for high achievement. The correlations with natural science achievement of .31 for computational interest, of −.33 for persuasive

interest, and of −.30 for social service interest are also significant and suggest plausible interpretations.

Achievement in the social studies test is, consistently enough, also moderately related to science interest. The correlation of .31 with literary interest is likewise consistent, but the negative correlation of −.37 with interest in social service is surprising. Scores on all of the achievement tests, however, are negatively correlated with social service interests, which may indicate that social service interests are inversely related to general academic ability.

The logically predicted association between English composition and literary interest is not satisfied by the correlation of .16 actually found. However, in this connection, it should be recalled that prescribed courses in both high school and college composition greatly limit free electives in this field, and therefore achievement is more likely to be motivated by extrinsic motives than by interest. The relatively high negative correlation of −.51 between English composition and clerical interest is perplexing; but since clerical interest also correlates slightly negatively with the other two achievement tests, it may mean that such interests are also inversely related to academic ability.

The fact that the distinctly highest interest-achievement relation found is in natural science should be related to the "exceptions" from the trend of low interest-achievement correlations reported in previous investigations and mentioned in the review cited in the introduction to this paper. Walter and Eurich reported correlations between interest and *general* achievement (Cooperative General Culture Test) for five different areas: between science-mathematics interest and general science achievement the correlation was .30. In the other areas they ranged from .05 to .19, (**15**, p. 570). Segel and Brintle found correlations between the Strong engineering key and the mathematics and science parts of the Iowa High School Content Examination of .49 and .36, respectively. For the medicine key, correlations with the same achievement tests were .28 and .29, respectively (**8**, p. 444). Triggs has also reported significant and logically consistent correlations between interest (Kuder's inventory) and general achievement as measured by the Iowa High School Content Examination as well as by the Iowa English Training Test: general science achievement and science interest correlated .32 and .42 for college men and women, respectively; for mathematics achievement and computational interest, the correlations were .34 and .36; and, contrary to the findings of this study, literary interest correlated .40 and .33 with English Literature and .30 (both men and women) with Iowa English Training (**12**). In all three of these instances general achievement tests were used. Turner, however, using grade-point-ratio but for the averages of

several courses within each field, also found significant correlations between science and mathematics achievement and interests. Kuder scientific interest correlated .29 with biological sciences and .32 with physical sciences; and Kuder computational interest correlated .49 with physical science achievement. In all other areas, however, where some relationship might logically have been expected, the correlations were zero or low.

In summary, the results of the present investigation together with the data from at least four other studies have demonstrated that science and mathematical interests are definitely related to general achievement in the parallel areas. For other areas, significant and logically consistent interest-achievement relationships have not been so clearly indicated, though some slight to moderate relationships have been noted for literature and for social studies. One interpretation of this outcome which suggests itself is that, except for the interest-achievement parallels in science and in mathematics the present tests of interests and achievements in the same fields are not psychologically concinnate.

2. *Relation of interests to election of courses in corresponding fields.* Implicit in Strong's hypothesis on the relation between interest and long-range achievement are the assumptions that the number of courses which students elect in a field should be significantly related to both the extent of interest in it and the amount of knowledge accumulated in the area. Data on this aspect of his hypothesis are presented in Table 2.

Table 2

Correlations between Per Cent of Natural Science and Social Studies Courses Taken in High School and College and Achievement and Interest. N = 81

Interests and Achievement	Per Cent of Total Credit in	
	Natural Science	Social Studies
Scientific Interest	.54	−.35
Social Service Interest	−.17	.32
Natural Science Achievement	.24	−.20
Social Studies Achievement	.08	−.07

As far as scientific interest and election of courses in natural sciences is concerned, Strong's hypothesis is very well confirmed. And, when all the factors known to affect achievement are considered, the relation between natural science achievement and number of courses taken in the field is also probably significant. It is also apparent that science interest leads to fewer elections of courses in social studies. The negative correlation of −.20 between natural science achievement and number of

courses taken in social studies probably means that the greater a student's achievement (and very likely interest) in natural science, the less likelihood of his electing courses in social studies. The relatively low correlations between social service interest and number of courses taken in natural sciences and in social studies indicate that students with higher social service interests tend slightly to avoid natural science and to elect social studies courses. It is puzzling to note that between social studies achievement and number of courses taken in social studies, there is not even a low positive relationship. A similar result, however, has also been reported by Learned and Wood (6). A suggestive interpretation of this apparent inconsistency is found in the facts that social service interest which correlates positively with number of courses taken in social studies appears to be negatively related to general academic ability. A disproportionate number of students with low academic ability may frequently elect social studies courses, and despite the relatively high number of credits earned in the field, because of their inferior abilities, they may do relatively poorly on a general achievement test in the field. This effect would tend to mask, in a correlation table, the otherwise expected relationship between amount of experience and knowledge in a field of study.

Conclusions

From the data on the intercorrelations among three variables—interests as measured by the Kuder Preference Record, general long-range achievement as evaluated by the USAFI Tests of General Educational Development, and the relative amounts of credit earned in the natural sciences and in the social studies—for 81 students who were about to complete their sixth quarter in college, the following *tentative* statements are made:

1. Strong's hypothesis that "Interests may not correlate to any great degree with achievement over a short period of time and yet may correlate significantly when achievement involves performance over a considerable period of time" is supported very well by the correlation of .50 between science interest and long-range achievement in natural science. The correlations with natural science of .31 for computational interest, of −.33 for persuasive interest, and of −.30 for social service interest are also significant and not inconsistent with the hypothesis.

2. The low but probably significant correlations between social studies achievement and interest in both science and literary activities are also consistent with the hypothesis.

3. The fact that social service interests correlate −.37 with social studies achievement appears to be inconsistent with the hypothesis; the

further fact that social service interest correlates negatively with all three achievement tests, however, indicates that this interest is inversely associated with general academic ability. The data also suggest that both persuasive interest and clerical interest are also slightly negatively related to general acadmeic ability.

4. The lack of correlation between literary interest and English composition would not have been predicted by the hypothesis; but the failure of the hypothesis in this instance may be due to the fact that English composition courses are largely prescribed in both high school and college and thus free electives in this field are very greatly restricted. This situation would likely result in achievement in English composition being motivated more by extrinsic motives than by intrinsic interest, and thus an otherwise possible relation between interest and achievement in this field would be masked. It was noted, however, that Triggs (12) did find relatively low but probably significant correlations between Kuder literary interest and general achievement in English literature.

5. The implication in Strong's hypothesis that interest in a given field would also lead to more frequent choice of courses in that area was confirmed by the data from both natural science and social studies, the correlations being .54 and .32, respectively, for the two fields of study.

6. The further expectation from the hypothesis that achievement on a general test in a field of knowledge would be related to the number of courses taken in that field was verified for natural science, but it was not confirmed for social studies.

7. The general trend of results of this study and the findings from related investigations would seem to justify the use of vocational interest inventories in guidance and also to support the otherwise well-established belief in the effectiveness of appeal to interest in teaching. Despite the fact that interests, because of the possible masking effect of extrinsic motives, have shown low or even zero correlations with achievement in single courses to which they seemed to correspond, they may for many students become abiding motives which serve to integrate learning experiences over long periods of time into effective patterns of knowledge.

Received February 5, 1946.

References

1. Allport, G. W. *Personality, a psychological interpretation.* New York: Henry Holt and Company, 1937.
2. American Council on Education. *Examiner's manual, tests of general educational development, college level.* New York: Cooperative Test Service, 1944.
3. Bolanovich, D. J., and Goodman, C. H. A study of the Kuder Preference Record. *Educ. & Psychol. Meas.*, 1944, **4**, 315–326.
4. Frandsen, A., and Darke, R. An evaluation of personality ratings and intelligence in predicting scholarship. *The Utah Educational Review*, 1935, **29**, 127–130.

5. Frandsen, A. Appraisal of interests in guidance. *J. educ. Res.*, 1945, **39**, 1–12.
6. Learned, W. S., and Wood, B. D. The student and his knowledge. *A report of the Carnegie Foundation on the results of high school and college examinations of 1928, 1930, and 1932.*
7. Pressey, S. L., and Robinson, F. P. *Psychology and the new education.* New York: Harper & Brothers, 1944.
8. Segel, D., and Brintle, S. L. The relation of occupational interest scores to achievement test results and college marks. *J. educ. Res.*, 1934, **27**, 442–445.
9. Shaffer, L. F. *The psychology of adjustment.* Boston: Houghton, Mifflin Company, 1936.
10. Strong, E. K. *Vocational interests of men and women.* Stanford University: Stanford University Press, 1943.
11. Stroud, J. B. *Psychology in education.* New York: Longmans, Green & Co., 1946.
12. Triggs, F. O. A study of the relation of Kuder Preference Record scores to other measures. *Educ. & Psychol. Meas.*, 1943, **3**, 341–354.
13. Turner, Alberta B. The predictive significance in home economics of tests of abilities, interests, adjustments, and study habits. *M.A. Thesis*, Logan, Utah: Utah State Agricultural College, 1941.
14. Turney, A. H. Intelligence, motivation, and achievement. *J. educ. Psychol.*, 1931, **22**, 426–434.
15. Walters, A., and Eurich, A. C. A quantitative study of the major interests of college students. *J. educ. Psychol.*, 1936, **27**, 561–571.

The Shipley-Hartford Retreat Scale as a Measure of Intellectual Impairment for Military Prisoners

Morse P. Manson, Captain, AGD and
Harry M. Grayson, Captain, AGD

The MTOUSA [1] Disciplinary Training Center is the largest overseas installation for the confinement of American prisoners who were adjudged guilty, by a General Courts Martial, of violating one or more of the Articles of War. As part of the psychological evaluation of all prisoners at the institution, a routine testing program was instituted.[2] The testing battery included: 1. The Qualification Test as used in induction stations for literacy; 2. The Shipley-Hartford Test for intelligence and mental deterioration; 3. The Cornell Selectee Index for psychoneurosis; and 4. The Minnesota Multiphasic Personality Inventory for the detection of serious personality disturbances, such as psychoneuroses and psychoses.

Those men who, by failing the Qualification Test, were illiterates according to Army standards, were given the following individual tests:

1. Army Individual Wechsler (seven sub-tests); and 2. Cornell Selectee Index (oral). Other tests given when necessary were: 1. Kent Emergency Test (intelligence); 2. Hunt-Minnesota Test for Organic Brain Damage; and 3. Visual Classification Test (Malingering Key). In addition, all Army test scores were taken from the WD AGO Form Number 20. Various test analyses, studies of inter-test relationships, and experimental investigations were carried out.

The present report is limited to an evaluation of the Shipley-Hartford Retreat Scale as applied to the General Courts Martial prisoner. The scale, which consists of a vocabulary test and an abstract-thinking test, was incorporated into the group testing program because it conveniently offered four significant measures: vocabulary age, abstract-thinking age, mental age, and conceptual quotient (CQ). The mental age is based on a weighted total (vocabulary plus abstract-thinking) score. The CQ is presumed to be a measure of intellectual impairment or deterioration. It is obtained by reference to a table of ratios between the abstract-thinking and vocabulary scores, ". . . based on the clinico-experimental observation that in mild degrees of mental deterioration, and in other

[1] Mediterranean Theatre of Operations, U. S. Army.

[2] An article describing the complete program, "The Psychological Clinic at the MTOUSA Disciplinary Training Center" appeared in the *American Psychologist*, Vol. I, No. 3, March, 1946.

conditions involving intellectual impairment, vocabulary is relatively unaffected, but the capacity for abstract-thinking declines rapidly."[3] Impairment is measured by the extent to which the individual's abstract-thinking falls short of his vocabulary.

In a prison situation, one of the advantages of an objective test measure of deterioration is that the examinee has not the slightest idea of what is being attempted. He is unaware of the fact that the *discrepancy* between the two test scores is the significant factor; and where he may attempt to malinger, the likelihood is that he will lower his scores on *both* parts of the test, thus minimizing the chances of undue discrepancy. In addition, such malingering will probably result in a vocabulary score less than 14.3 years, below which according to the test manual, the test is not considered a valid measure of deterioration.

Table 1

Per Cent of Norm Group Falling Below Designated CQ's and Claimed Clinical Significance

CQ	Per Cent of Group	CQ	Clinical Significance
90 and below	27	Above 90	normal
85 and below	17	85–90	slightly suspicious
80 and below	10	80–85	moderately suspicious
75 and below	5	75–80	quite suspicious
70 and below	2	70–75	very suspicious
65 and below	1	Below 70	probably pathological

The frequencies with which various degrees of disparity were encountered in the standardization group of 1,046 normals, as indicated by sub-average CQ's and the interpretative ranges for the CQ, are given in Table 1.

Prisoner CQ Percentages

During the course of working with the test, it became apparent that CQ's supposedly indicative of various degrees of deterioration were occurring with considerable frequency (See Table 2). An examination of the CQ scores of 1262 prisoners who had earned vocabulary age scores above 14.3 was undertaken. The statistics, compared with the statistics of the standardization group, showed very marked differences.

According to the standardization figures, about three quarters of the group had CQ scores indicating complete absence of impairment, whereas

[3] *Shipley-Hartford Retreat Scale for Measuring Intellectual Impairment: Manual of Directions and Scoring Key.* Neuro-Psychiatric Institute of the Hartford Retreat, Hartford, Connecticut, 1940.

much less than half of the prison population showed such scores. Fully 14 per cent of the white and 17 per cent of the negro group had scores interpretable as "probably pathological deterioration." The average CQ score for the prison group was 88.8, which, according to Table 1, indicates some "slight suspicion" of deterioration for the group as a whole. The evidence thus far would appear to indicate that: (1) either the prisoner population contains an extremely high percentage of men suffering from intellectual impairment; or (2) the norms are in need of revision, at least for this type of group. Such revision could be easily effected by keeping the Shipley norm percentages, and equating them against the actual CQ's obtained on the prison population. For example, 24 per cent of the total prisoner group receive CQ's of 75 and below, interpretable as "very suspicious," whereas, according to the Shipley norms, a similar percentage

Table 2

Percentage Distribution of CQ's of Prison Group in Comparison with Norm Group

CQ	Shipley Norms (N 1046)	Prison Group		
		White (N 1138)	Negro (N 124)	Total (N 1262)
90 and below	27%	59%	61%	59.1%
85 and below	17	45	52	44.8
80 and below	10	34	41	34.6
75 and below	5	24	25	24.0
70 and below	2	14	17	14.2
65 and below	1	7	6	6.9

(by interpolation) receives CQ's of 88.5, interpretable as "slightly suspicious." By making the interpretations identical for similar CQ percentages, prisoner CQ's between 70 and 75 would be interpreted as "slightly suspicious" rather than as "very suspicious."

The high percentage of CQ's indicative of some degree of impairment, which was obtained on the prison groups, pointed to the fact that the abstract-thinking age scores should be considerably below the vocabulary age scores for this group. A distribution of vocabulary ages, abstract ages, and mental ages of 470 prisoners who ran the entire gamut of Shipley scores (not restricted to vocabulary ages of 14.3 or more), was drawn up, and a difference of 1.42 years was found between the vocabulary and abstract-thinking scores. Vocabulary Age = 14.18 years; Abstract Age = 12.76 years; Mental Age = 13.28 years. This difference would probably have been even greater if the cases containing vocabulary age scores below 14.3 were eliminated.

The question is raised as to whether the prisoner CQ scores may be unduly low either because the abstract-thinking age scores are spuriously low, or because the vocabulary age scores are spuriously high.

The Shipley Test and the AGCT

Examination of the various Army and Disciplinary Training Center test scores suggested that it was the vocabulary age scores which were out of line. For example, the average AGCT score for the entire prisoner population was only 84.7, compared with an Army average score of 100; whereas the Shipley vocabulary age was at the average adult level, which would therefore correspond to the average Army AGCT score. The Shipley test, however, was given only to prisoners who were literate (scored 9 or more) on the Army Qualification test. Accordingly, the average AGCT score for literates was computed, and was found to be 90.7 for the white group and 82.7 for the negro group. In order to make

Table 3

Mental Age, IQ, and AGCT Equivalents

Mental Age	AGCT & IQ	Mental Age	AGCT & IQ
18–6	130	13–6	95
18–0	126	13–0	91
17–6	123	12–6	88
17–0	119	12–0	84
16–6	116	11–6	81
16–0	112	11–0	77
15–6	109	10–6	74
15–0	105	10–0	70
14–6	102	9–6	67
14–0	98	9–0	63

the comparison of AGCT and Shipley scores more meaningful, Table 3 showing mental age equivalents for AGCT scores was worked out. This table is actually a table of mental age and IQ equivalents, but is applicable to AGCT scores since the average is 100 in both cases and the respective standard deviations of 17 and 20 are not too dissimilar. It was constructed by making an IQ of 70 equal to a mental age of 10.0 years, since these figures are frequently used interchangeably to establish a line of demarcation for mental deficiency. The average adult age then becomes 14.3 years. There is evidence in the literature that this is a fairly good approximation to average adult intelligence, and produces most constant IQ scores.[4]

[4] Pintner, Rudolf. *Intelligence testing.* New York: Henry Holt & Co., 1932.

The Shipley and AGCT test scores, derived from literates, could now be directly compared as follows:

	Age Equivalent
Shipley Vocabulary	14–2
Shipley Abstract-thinking	12–8
Shipley Mental Age	13–3
AGCT (white)	12–11
AGCT (negro)	11–10

It is apparent that the Shipley vocabulary-age score is higher than the others. (It is also responsible for raising the Shipley Mental age score into which it enters.) It was felt, therefore, that the AGCT might be a more valid measure of original intellectual level than the Shipley vocabulary test as a base against which to measure deterioration. A distribution of 100 cases, selected at random from among those men who had earned Shipley vocabulary age scores of 14.3 or more, was drawn up. The average AGCT score for this group was 100.4, or approximately 14.3

Table 4

Percentages Falling Below Designated CQ's Derived from Shipley and AGCT

CQ	Shipley Form CQ's	AGCT-CQ's
90 and below	27%	29%
85 and below	17	19
80 and below	10	13
75 and below	5	6
70 and below	2	4
65 and below	1	2

years. (Note that all these cases were not only literate, but had earned at least minimal Shipley vocabulary scores for valid use of the test.) The average CQ for this group was 88.5, or 0.3 points less than the average CQ score cited previously based on 1262 cases. The average AGCT-CQ (CQ's obtained by substituting AGCT mental age equivalents for the Shipley verbal age scores) was 98.9. It is obvious that the AGCT-CQ shows far less impairment than does the Shipley CQ. The percentages showing scores indicative of deterioration based on the AGCT-CQ are now very similar to the percentages given for the original standardization group as may be seen in Table 4.

The AGCT had been administered upon entrance into the Army. For most prisoners this had been three years prior to their examination at the Disciplinary Training Center. Whether the group, most of whom had been subjected to combat and other tensions, actually had undergone

impairment to the extent indicated by the Shipley CQ, or whether the vocabulary test performance was unduly high for the prisoners, thus making the AGCT a better base against which to measure deterioration, remains an open question. Unfortunately, re-test AGCT's were not administered, so that possible changes in AGCT score were not available for comparison. It is the opinion of the writers that the Shipley vocabulary test scores are unduly inflated for the prisoner group, and that the AGCT is a more accurate base measure of original intellectual capacity. This does not mean that the AGCT need be substituted for the Shipley vocabulary test, but it does suggest a downward revision of the vocabulary norms, possibly by equating them against the AGCT mental-age equivalents for the prisoner population.

Some Factors influencing the CQ

From the viewpoint of multiple causation, it is theoretically possible for factors other than deterioration to yield "deterioration" scores. For example, it is conceivable that simple misunderstanding of the directions on the abstract-thinking test may yield a significant differential. It is also possible that unfamiliarity with the abstract test, coupled with insecurity in a new situation, might result in emotional blocking and con-

Table 5

Numerical Distribution of Shipley CQ's and AGCT-CQ's (N = 100)

	Shipley CQ	AGCT-CQ
140–144	—	1
135–139	—	2
130–134	—	—
125–129	—	2
120–124	1	6
115–119	2	7
110–114	4	4
105–109	11	14
100–104	8	14
95– 99	7	10
90– 94	10	11
85– 89	18	10
80– 84	10	6
75– 79	8	7
70– 74	12	2
65– 69	6	2
60– 64	2	2
55– 59	1	—
Total	100	10

sequent penalization (temporary sub-typical performance) on this phase of the test which requires mental agility. Inadequate motivation, especially in a prisoner group, and consequent lack of effort on the abstract-thinking test, which requires a higher degree of concentration, may yield undue discrepancy. Fatigue, resulting from the extremely rigorous field training program, may also lower effort more on the abstract-thinking, which calls for fill-in responses, than it does on the vocabulary test, which calls for multiple-choice responses. Also, many of the obtained discrepancies could probably be explained as genuine differences in the actual original ability on these different aspects of intelligence, since different mental functions are tapped, and different type responses are required.

As a matter of fact, if the CQ's are distributed in a normal fashion, with just about as many earning above-normal CQ's as earn below-normal CQ's then "chance factors" could explain both parts of the distribution equally well, making superfluous the concept of deterioration as an explanation of the lower half of the curve.

The distribution of CQ's given in Table 5 do follow such a normal curve. The data are given for the Shipley CQ's and for the AGCT-CQ's, each based on the same 100 cases.

"Individual Idiosyncrasy"

At this point the question arises as to how much of a difference in either direction can be ascribed to "individual idiosyncrasy" favoring one or the other test function. In order to determine the extent of significant individual idiosyncrasy, it is necessary to determine the standard error of estimate of a score; and this, in turn, requires computation of the reliability of the vocabulary and abstract tests, and their intercorrelation. In line with Kelley's treatment,[5] the essential data based on 100 cases selected at random, are shown in Table 6.

[5] The data are based on the following formulae:

$$r_{vv} = \frac{2r_{11}(v)}{1 + r_{11}} = \frac{2(.70)}{1.70} = .82$$

$$r_{aa} = \frac{2r_{11}(a)}{1 + r_{11}} = \frac{2(.85)}{1.85} = .92$$

$$\sigma_{v_\infty} = \sigma_v\sqrt{1 - r_{vv}} = 5.28\sqrt{1 - .82} = 2.22$$

$$\sigma_{a_\infty} = \sigma_a\sqrt{1 - r_{aa}} = 8.30\sqrt{1 - .92} = 0.66$$

$$r_{v_\infty a_\infty} = \frac{r_{va}}{\sqrt{r_{vv}}\sqrt{r_{aa}}} = \frac{.40}{\sqrt{.82}\sqrt{.92}} = .70$$

$$\sigma_{D_\infty} = \sqrt{\sigma^2_{v_\infty} + \sigma^2_{a_\infty} - 2r_{v_\infty a_\infty}\sigma_{v_\infty}\sigma_a}$$

$$= \sqrt{(2.22)^2 + (.66)^2 - 2(.40)(2.22)(.66)} = 2.18$$

Table 6

Correlations and Standard Deviations for the Vocabulary and Abstract Thinking Tests

	Self-Correlation	Spearman-Brown Correction	Raw R-L Correlation	Attenuated R-L Correlation
Correlations				
Vocabulary	$r_{11}(V) = .70$	$r_{vv} = .82$	$r_{va} = .40$	$r_{v_\infty a_\infty} = .70$
Abstract	$r_{11}(A) = .85$	$r_{aa} = .92$		
Standard Deviations				
	Raw S.D.	Attenuated S.D.		Attenuated S.E. Diff.
Vocabulary	$\sigma_v = 5.28$	$\sigma_{v_\infty} = 2.22$		$\sigma_{D_\infty} = 2.18$
Abstract	$\sigma_a = 8.30$	$\sigma_{a_\infty} = 0.66$		

The logic underlying the statistical treatment follows: If the vocabulary and abstract thinking scores were perfectly correlated, any difference at all between them would be statistically significant. The lower the correlation, the greater must be the difference in order to establish significance, since the error of estimate of a predicted score varies inversely with the size of the correlation coefficient.

Table 7

Distribution of Differences in Favor of Vocabulary or Abstract Thinking Tests

Individual Differences	Vocabulary Superior	Abstract-Thinking Superior
7.5–7.9	1	—
7.0–7.4	—	—
6.5–6.9	1	—
6.0–6.4	2	—
5.5–5.9	3	—
5.0–5.4	2	—
4.5–4.9	3	2
4.0–4.4	5	—
3.5–3.9	5	—
3.0–3.4	15	1
2.5–2.9	5	4
2.0–2.4	2	—
1.5–1.9	10	6
1.0–1.4	10	1
0.5–0.9	8	6
0.0–0.4	5	3
Total	77	23

The attenuated standard error of the difference between the vocabulary age and abstract age scores being 2.18 years, it follows that individual differences in either direction must be greater than 3×2.18, or 6.54 years, to be statistically significant. The distributions of differences favoring the vocabulary and the abstract thinking tests, respectively, appear in Table 7.

Only two cases show significant "individual idiosyncrasy" (differences greater than 6.54 years) and both of these favor the vocabulary test. The remaining differences in either direction are all explicable in terms of expected vocabulary-abstract differentials (CQ's). In other words, the standard error of estimate of test scores, derived from the reliability, variablity, and intercorrelations of the vocabulary and abstract thinking tests, adequately accounts for the obtained differences, and makes gratuitous the interpretation of the CQ as a measure of "deterioration," at least for the prisoner population examined. This is in decided contrast to the extent of deterioration indicated for this group according to the original Shipley norms and table of interpretation.

Reliability

The test has a good degree of reliability, the coefficients being .82 for the vocabulary and .92 for the abstract thinking test. It is interesting that the abstract thinking test, although containing only twenty items, is so highly reliable, even more so than the forty item vocabulary test. The greater reliability is explained largely by the fact that the shorter abstract thinking test in actual fact produces greater variablity of test scores than the vocabulary test, as shown by their respective standard deviations. As a result, the abstract thinking test also has greater weight than the vocabulary test in determining the total mental age score. The correlation between the tests and Shipley mental age are as follows:

(Vocabulary and mental age) = .81
(Abstract thinking and mental age) = .91

Of course, both correlations are spuriously high, since the mental age scores are made up of the vocabulary and abstract thinking test scores.

Consistency

As a further check on the reliability and consistency of its various measures, the Shipley test was readministered to a group of 300 cases selected at random. The time interval between the two tests varied for the different prisoners, who had received the first test at different times, but the length of time between tests probably averaged several months. For this group of 300 cases the following reliability coefficients were

obtained:

r (Vocabulary) = .78
r (Abstract-thinking) = .84
r (Mental age) = .85

Those retest coefficients, although lower than the odd-even coefficients, still indicate a fairly high degree of reliability. The average CQ for 172 cases with vocabulary ages of 14.3 and above was 86.2 on the first test and 89.1 on the second. These are in fairly close agreement with each other, and with the average CQ of 88.8 obtained from the original test scores for 1262 cases. They also fall within the same CQ interpretation range in the Shipley table.

The Mental Age Score

The relatively high reliability of the short and easily administered Shipley test might well make it a good measure of intelligence. In order to check this point, the Shipley mental age scores and the AGCT scores were correlated on a sampling of 100 cases, producing a coefficient of .75. In addition, the average Shipley mental age was found to be 13 years 3 months and the average AGCT mental age equivalent was found to be 12 years 11 months. This difference of 4 months in average mental age is probably not sufficiently large to require correction. The correlation between the Shipley mental age and the AGCT mental age equivalent appears sufficiently high to warrant the use of the Shipley as a group measure of intelligence.[6]

CQ's of Different Groups

In order to determine whether there were any differences in CQ for various types of groups among the prisoners, the CQ distributions of the following groups were drawn up:

Groups	Characterization
NSPD :	No significant personality defects
MSPD:	Mild personality defects
CPS :	Constitutional psychopath
CPSI :	Inadequate personality
CAL :	Chronic alcoholic
ALT :	Alcoholic tendencies
PNS :	Psychoneurosis, severe
PNM :	Psychoneurosis, mild

[6] In actual practice at the Disciplinary Training Center, the Shipley test proved very useful as a measure of intelligence. The Shipley mental age, if 10 years or more, was used in the case of literates for whom no AGCT score was available. It was also used where the AGCT score appeared unduly low, compared with the amount of education. All illiterates and prisoners with Shipley mental ages less than 10 years and AGCT scores below 70 were given the Army Wechsler Individual Test.

These groups were so classified jointly by the psychological and neuropsychiatric departments at the institution. The classifications were based on very extensive case reports covering hundreds of items dealing with family background and stability, educational and occupational adjustment, marital-sexual relationships, habits and traits of character, health, emotional stability, military history, and institutional adjustment.

The percentage distributions of CQ's for cases achieving vocabulary age scores of 14.3 or more, the average and CQ's for these groups, are presented in Table 8.

Unfortunately the number of cases in the various groups is relatively small, ranging from 16 to 100 cases. Seven of the eight white groups have at least 50 cases, whereas only one of the negro groups has more than 42 cases. Nevertheless, the average CQ's of white and negro prisoners for these various groups are remarkably similar, the differences ranging from 2 points (for four groups) to 6 points for one group. However, the percentages falling at the various CQ intervals, indicative of various degrees of suspicion of deterioration, show relatively marked differences for the white and negro groups. The minimum percentage for any group falling at "90 and below," indicative of slight suspicion of impairment, is 52 per cent as against a corresponding 27 per cent on the Shipley norms. Using the CQ of 70 and below (indicative of "probably pathological deterioration") the percentages range from 7 for the NSPD group to 25 for the negro CPS group. If groups containing less than 35 cases are ignored, the corresponding percentages range from 7 for the NSPD group to 22 for the white CAL group. This white chronic alcoholic group shows the greatest spread of CQ's, with fully nine per cent showing scores of "60 and below." It seems likely that for this group, "deterioration" of the central nervous system, brought about by excessive drinking, may be the true explanation. Another observation of interest is the high percentage of white PNS cases falling at "90 and below." Sixteen of the twenty cases (80%) are included here, giving some support to the idea that emotional blocking may be a factor in the production of vocabulary-abstract thinking differentials. Based on the average CQ, the group showing the greatest amount of discrepancy between vocabulary and abstract-thinking ability is the CPS group, both for white and negro prisoners. The average CQ's, 76 and 78, respectively are interpreted by the Shipley table as "quite suspicious" of deterioration. This marked discrepancy for a group of constitutional psychopaths appears to be of genuine significance, but whether it indicates deterioration, or original defectiveness in abstract-thinking (and consequent asocial behavior), or something else, is still a matter for conjecture.

In order to check on whether defective abstract-thinking is characteristic of the constitutional psychopathic group, the average vocabu-

Table 8

Percentage Distributions of CQ's for Clinically Diagnosed Groups

CQ	Shipley	NSPD		MSPD		CPS		CPSI		CAL		ALT		PNS		PNM	
		W	N*	W	N	W	N	W	N	W	N	W	N	W	N	W	N
90 and below	27	59	55	52	59	56	88	57	64	58	64	60	70	80	66	56	67
85 and below	17	45	39	36	49	37	88	39	57	47	59	38	70	70	56	44	62
80 and below	10	39	26	25	41	31	63	27	48	40	45	28	57	45	41	34	49
75 and below	5	31	19	17	27	30	31	18	36	38	27	23	37	15	31	22	31
70 and below	2	18	16	7	19	19	25	14	21	22	22	18	23	10	18	13	21
65 and below	1	12	6	4	8	9	6	4	2	8	9	7	3	0	9	5	18
60 and below	0	2	0	3	5	2	0	0	0	7	0	3	3	0	4	3	5
55 and below	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	1	0
Average CQ		85	87	89	85	76	78	87	85	83	87	85	81	82	80	87	81
No. of Cases		51	31	69	37	54	16	51	42	60	22	68	30	20	100	100	39

* W and N stand for white and negro, respectively.

Table 9

Mean Vocabulary Age and Mean Abstract Thinking Age for Clinically Diagnosed Groups

	Total Group								Partial Group (VA = 14.3 or more)							
	White				Negro				White				Negro			
Group	No.	VA	AA	D*	No.	VA	AA	D	No.	VA	AA	D	No.	VA	AA	D
NSPD	100	14.5	12.9	1.6	72	13.7	13.7	0.0	53	16.2	13.8	2.4	30	16.8	13.9	2.9
MSPD	100	14.6	13.4	1.2	94	13.4	12.2	1.2	61	15.9	14.3	1.6	37	15.8	13.7	2.1
CPS	97	14.1	13.0	1.1	33	13.5	11.5	2.0	57	16.3	13.9	2.4	15	16.2	12.8	3.4
CPSI	97	14.6	13.1	1.5	100	13.8	12.7	1.1	53	16.5	13.8	2.7	47	15.9	13.5	2.4
CAL	86	14.8	12.7	2.1	42	13.7	11.9	1.8	51	16.5	13.7	2.8	21	14.2	12.9	1.3
ALT	98	15.7	14.2	1.5	98	13.5	11.7	1.8	73	16.6	14.0	2.6	37	16.5	13.2	3.3
PNS	79	14.5	12.4	2.1	45	13.6	11.7	1.9	47	16.4	13.5	2.9	18	16.1	13.4	2.7
PNM	100	14.3	12.8	1.5	95	13.5	11.8	1.7	58	15.9	13.9	2.0	37	15.9	13.2	1.7
Mds	98	14.6	12.9	1.5	83	13.6	11.9	1.7	55	16.4	13.9	2.5	34	16.0	13.3	2.7

* D = VA minus AA.

lary and abstract ages for the various groups were computed. This was done twice, once for all individuals within each group, and once for individuals with vocabulary age scores of 14.3 or more. The data appear in Table 9.

As is to be expected, the vocabulary age scores of the partial group (earning vocabulary ages of 14.3 years or more) are considerably above those of the total group. This difference carries over, but not to the same degree, on the abstract thinking age scores.

The supposition that defective abstract thinking ability is responsible for psychopathic behavior does not hold up. For the total white group the average CPS abstract thinking age score is above the median for all groups, including the NSPD group. For the partial white group, it is exactly at the median for all groups. For the negro groups, the average CPS abstract thinking score is the lowest score for all groups but it is not sufficiently low, compared with the others, to single out the negro constitutional psychopath as especially defective in abstract thinking.

Although practically all the average differences between the vocabulary and abstract scores would appear to be statistically reliable, the *comparative* differences for the various breakdown groups can most probably be attributed to sampling errors, especially in view of the small number of cases in many of the groups. For example, a difference of 3.3 years for the partial ALT negro group might appear significant, except that it is greater than the difference for the extreme CAL group, the latter difference being less than the median difference for all groups. The conclusion appears to be that, although there are differences in the vocabulary-abstract differentials of various breakdown groups, they are not sufficiently marked, compared with each other, to establish statistical reliablity.

Summary and Conclusions

Various statistical analyses, following assorted experimental manipulation, were performed on the Shipley Hartford Retreat Scale, with the fundamental objective of ascertaining its validity as a measure of intellectual impairment or deterioration for General Court prisoners. Incidentally, data relative to test reliability, as well as validity as an intelligence measure, emerged.

It was found that, if the CQ interpretation table were taken at face value, the majority of prisoners evidenced some degree of deterioration. The preponderance of scores indicating impairment raised the question of the validity of the instrument for a military prisoner group. Either the prisoner population as a whole had undergone deterioration, or the obtained vocabulary-abstract differentials could not be interpreted fairly against the Shipley norms. Examination of the average test scores on various measures, especially the AGCT, made it appear that the obtained

vocabulary age scores were unduly high for the prisoner group. Substitution of the AGCT for the Shipley vocabulary test as a base against which to measure deterioration produced far less startling results, which were in much closer agreement with the original Shipley norms. However, the distribution of CQ scores conformed to the normal probability curve; and especially for the AGCT-CQ's, the curve showed as many scores above 100 as below. If the above-normal scores could be attributed to chance, why introduce the concept of "deterioration" to explain the below-normal scores? This view was further buttreessd by an investigation into "individual idiosyncrasy" favoring either test, which revealed that all but two per cent of the test scores fell within the limits of the standard error of estimate of the predicted scores on either test, and were thus explainable entirely on the basis of chance.

The odd-even and retest reliability coefficients of the vocabulary and abstract-thinking tests were found to be quite high, with the abstract thinking test giving the higher reliability coefficient and contributing more to the total mental age score. The average CQ of initial and retest administrations showed a high degree of consistency. By correlating and comparing the Shipley mental age scores with the AGCT mental age equivalents, it was found that the relationship was sufficiently high to justify the use of the Shipley as a group test of intelligence.

An examination of the CQ's for various types of groups, including normals, and moderate and severe psychopaths, alcoholics, and psychoneurotics, revealed the possibility of some group differences, and suggested that different factors might account for the CQ distributions of the different groups. For example, the disproportionate percentage of white chronic alcoholics earning extremely low CQ's, well below the limit interpreted as "probably pathological deterioration," suggests the possibility of actual deterioration of the central nervous system as a result of excessive drinking. The high percentage of white severe psychoneurotics falling at or below the "slight suspicion" level suggests that emotional blocking, in a novel test situation calling for mental agility, may be responsible for producing vocabulary-abstract thinking differentials. However, all of these interpretations are weakened by the fact that most discrepancies, at least for the group as a whole, can be explained away largely on the basis of chance differences.

For a General military prisoner group, it appears that the norms for the interpretation of CQ's are in need of revision, or that the vocabulary age scores should be reduced. Further experimental research is necessary in order to identify, and determine the influence of, various factors which may affect the CQ. At present, the Shipley-Hartford Retreat Scale is a valid measure of intelligence, but is not a valid measure of deterioration, for the literate General Court prisoner.

Received January 4, 1946.

The Special Leadership Training Program at the Signal Corps Officer Candidate School *

Robert E. Taubman

The Signal Corps Officer Candidate School

Throughout the war years, considerable interest has been centered upon the problem of "leadership." This interest has not diminished with the cessation of hostilities. The army has had to face at every turn the question of what constitutes leadership, how to measure the qualities subsumed under this term, and how to develop these attributes. The various officer candidate schools within the several branches and services of the army have manifested special concern over the question of leadership. One school, the Signal Corps Officer Candidate School, Fort Monmouth, N. J., faced with approximately 25% of its washouts or failures through lack of leadership, has attempted a program to reclaim those candidates who, except for their undeveloped qualities of leadership, might otherwise prove excellent officers. This report is based on the program and results of the Special Training Detachment of the Signal Corps Officer Candidate School.[1]

This particular OCS was activated in July 1941 to meet the projected needs of our expanding army for signal officers. Its class enrollments have varied with the needs of the army throughout the course of the war. The first class started with some 490 candidates and succeeding classes increased in size to approximately 2,000 candidates by the following year and then decreased quite irregularly until class enrollment reached about 100 candidates per class.

Originally, OCS was essentially an academic course lasting three months. Later a field problem of about three days was added and still later the course consisted of three months of academic (basic army and Signal Corps) subjects *plus* approximately one month of training under simulated field conditions. Approximately 27,000 specially-selected enlisted men have attended the Signal Corps OCS since its inception, and of this number about 21,000 commissions have been granted to candidates who completed successfully all the course requirements. At the same time, as compared with the number of successful students, some

* Appreciation is extended to Lt. Col. W. F. Starr, Major J. H. Bair, Captain G. C. Steinbach and S.Sgt. W. M. Eastman, all of the Signal Corps OCS, for providing the data used herein and for suggestions regarding this paper.

[1] Since the writing of this report, the Special Training Detachment has been deactivated.

6,000 candidates have been eliminated for one or more indicated reasons. Table 1 presents a brief summary of the number and classification of failures at this OCS.

As can be seen from Table 1, the main causes for elimination at the school have been due to academic failure. Those relieved for purely academic reasons constitute 48.30% of all the washouts at the Signal Corps OCS. Similarly, and second in importance, is the great number of

Table 1

Number and Classification of Failure at the Signal Corps OCS *

Classification	Number of Failures	Percentage of Total Failures
Academic Failure	2,972	48.30
Lack of Leadership	1,501	24.39
Unsuccessful Turn Back	682	11.08
Lack of Technical Experience	383	6.22
Physical Reasons	206	3.35
Unbecoming Conduct	168	2.73
Lack of Adaptability	113	1.84
Miscellaneous	133	2.16

* Notes on Table 1: Explanation of the various categories.

Academic Failure: includes those candidates who have demonstrated inability to handle the basic army and Signal Corps subjects by excessive deficiencies in the more important courses, such as map reading, army organization, tactics and techniques of signal communication, principles of wire and radio, etc. These candidates are relieved through action of the Commandant of OCS upon recommendation of the Faculty Board.

Unsuccessful Turn Back: turn back students are those who, for a variety of reasons, were made to repeat any of the academic or field phases of the OCS curriculum. For the classes studied, 2,539 candidates were turned back, of whom 1857 or 73.14% were commissioned after successful make-up of deficiencies. The data at hand do not permit an analysis of the 682 failures listed.

Lack of Leadership: consists of candidates who, according to reviewing boards of officers, are so lacking in leadership qualities that special leadership training would be useless.

Lack of Technical Experience: this category comprises students who cannot cope with the special courses and practical work along Signal Corps lines.

Physical Reasons: inability to meet the final type of physical examination required for officers and certified disability for discharge (CDD) under regulations established by the Office of the Surgeon General.

Unbecoming Conduct: includes offenses under the generic term of conduct unbecoming an officer candidate, such as AWOL, misbehavior, infraction of special school rules, cheating, etc., as determined through action of the Commandant of OCS upon the recommendation of the Faculty Board.

Lack of Adaptability: extreme inability to adjust to the rigorous program at OCS.

Miscellaneous: discharge for reasons other than medical, including resignations, dependency, etc.

candidates who are eliminated because of their apparent lack of those leadership qualities deemed necessary in an officer. This accounts for 24.39%. Those who are unsuccessful after turn back make up 11.08% of all washouts. Lack of technical experience causes 06.22% of the attrition at the school, unbecoming conduct causes 02.73%, physical reasons 03.35%, lack of adaptability 01.84%, and the various minor groupings under miscellaneous causes of relief 02.16%.

It is not our intention here to deal with all the major categories of relief from the Signal Corps OCS. So far as the large number of academic failures are concerned, it should be pointed out that the school has provided every opportunity for success in this respect. The very idea of the Turn Back is an indication that the school is willing to give the candidates who are academically deficient a "second chance." Special evening classes are held, the candidates are required to attend officer-supervised "study hall" every evening for two hours, and the officer-instructors hold special review classes after the required study hall periods are over. Perhaps the large number of academic failures might be reduced by more effective screening at the posts, camps or stations from which the candidates are sent to OCS.

Our primary concern here is the 25% who are relieved for lack of leadership. It was the opinion of the OCS administration that a good part of the number of candidates relieved under this category suffered mostly because of lack of experience in handling men in the army situation, in giving them clear and concise orders and in the supervision of their work. Hence, commencing with Class 21 there was begun a system of providing leadership experience for those men who might otherwise have been included under the figure of relieved for lack of leadership.

Those who were turned back for leadership were farmed out, as it were, to various non-OCS training companies on the Post for the purpose of drilling troops, instructing them in basic military subjects, leading the men on hikes, etc. While 66% of the men thus turned back eventually graduated, it was felt that this method of training in leadership suffered from several standpoints. For one, adequate supervision of the candidates was impossible. Secondly, it soon became apparent that the turned back candidates were developing into good non-commissioned officers, rather than developing as potential officers. The company commanders to whom the men had been attached had not given them the right kind and sufficient amount of responsibility within the training unit.

The Special Training Detachment

Beginning with Class 38, a new plan was designed for providing leadership training. Under the guidance of Lt. Col. W. F. Starr, Assistant

Commandant of OCS, the Special Training Detachment was set up at Lincroft, N. J., about 6 miles from Fort Monmouth proper.

The detachment at Lincroft was based on a small, field unit pattern. Its physical plant consisted of 6 large pyramidal tents for sleeping quarters, a headquarters tent, commissioned officers' tent, two supply tents, a semi-permanent school building (1 room) and field sanitation installations. A large level stretch of ground served as combination drill field and athletic grounds. A novel feature was a specially-constructed drill obstacle course designed to test the candidate's skill in maneuvering a squad of men in a set pattern requiring a given number of commands with the proper time sequences. A detailed description of the installation is unnecessary. It is sufficient to point out that the Special Training Detachment was designed and functioned as an independent unit operating out of Fort Monmouth.

The basic philosophy underlying the detachment was: (1) to provide a training area with sufficient facilities for a small unit of about 50 candidates; (2) to construct a situation where candidates could assume the responsibilities and gain experience in running a model army unit; (3) to provide an atmosphere freed from the tension and anxieties of OCS where candidates could live and work as a group; (4) to provide opportunities for guidance in developing the attributes of leadership.

Candidates were referred to the Special Training Detachment in either of the two phases of OCS training: at the end of academic training or after the field exercise. In addition to these candidates, the school referred to the detachment pre-OCS men who had reported early from their original post, camp or station and had to wait for their assigned class to begin. This applied especially to many candidates who had come from overseas installations and frequently had to wait from 2 to 4 weeks for their class to begin OCS proper. Hence the detachment size varied from 10 to 50 men during the training cycle, which lasted 4 weeks.

The detachment was run completely, subject to supervision by the attached commissioned officers, by a rotating group of candidate officers. In the scheme of company administration were: Candidate Detachment Commander, Candidate First Sergeant, Candidate Plans and Training Officer, Candidate Supply Officer. These men held their positions for a period of 3 days, after which time a new set of company officers was appointed. The remaining candidates turned back for leadership training and the pre-OCS men constituted the detachment troops.

Each of the various official detachment positions had responsibilities similar to those faced by company officers in the usual army unit. To understand the particular training that the Special Training Detachment provided, it is necessary to examine the structure of these positions held by the candidates.

The Candidate Detachment Commander is faced with the overall supervision of the detachment. In the army, it is a familiar catechism that all things are a function of command. The company commander is charged with the efficient operation of his entire command, and this includes the proper organization and operation of the training program, maintenance of healthful living conditions, and the general welfare of his troops. It is his job to supervise the entire installation to see that the mission of the unit is fulfilled. He is charged with the adequate administration of all company records, property and funds. He is the coordinator of all subordinate officers and men. His position, as chief executive of the unit, involves him, in a supervisory capacity, in all the activities of the unit.

The Candidate First Sergeant is, as he is in any company, the commanding officer's chief executive secretary. He prepares most of the company administrative reports and especially the morning report which is the unit's daily history and appoints details of men for various company duties (guard, cleaning details, etc.). He is the liaison between the company commander and his men and carries out all orders of the commanding officer. No unit can operate successfully without an efficient first sergeant.

Two Candidate Plans and Training Officers are appointed for a given week. It is their task to prepare a training program for this one-week period. This schedule, in addition to fulfilling the requirements of certain army training programs, might include a number of exercises patterned after the main field problem to which the candidates are sent after the month's special leadership training. Thus, the Plans and Training Officers are encouraged to use their ingenuity in preparing training schedules. They have to prepare a balanced program, choose references for the various topics, arrange with Supply for the use of training aids, appoint instructors for the classes and are responsible for the smooth running of the training program.

The Candidate Supply Officer is in charge of all the property assigned to the detachment. His responsibility covers a wide range of materials necessary for the operation of the unit. In addition, he has to arrange with the supply section of OCS at Fort Monmouth for any new supplies that are needed for the detachment, and in so far as the training program is concerned, he has considerable work in the requisitioning of new materials. For the period of his tenure, the Candidate Supply Officer has extensive experience in checking equipment, signing out for highly valuable supplies, maintaining the unit's property in operating condition, and in the involved administrative forms which are part of army supply procedures.

In outlining briefly the officer positions held by the candidates, our main purpose is to indicate that the candidates are actually acquiring work experience in positions similar to those that they might have held in any usual army organization. They are expected to know the technical details of their job and especially to organize their work for the successful operation of the detachment in its various phases. All of these positions require the use of executive ability, work organization, planning and extensive contact with all the men in the detachment. And it is the belief, supported by obtained data, of OCS officials, that it is these types of experience which provide the basis for the development and maturation of the qualities of leadership necessary in an army officer. This is the crux of the program at the Special Training Detachment.

In addition to the above positions, all turned back candidates are required to instruct or take charge of the classes set up in the weekly training schedule. Usually, this amounts to two or three hours per week per candidate. Hence, opportunity is provided for experience in one of the most important roles that an officer fills—army teacher. Candidates are expected to demonstrate that they can prepare and organize an adequate class in any of the basic army and Signal Corps subjects called for in the training program. All classes are supervised by the attached commissioned officers and the candidate instructors observed by these officers.

All jobs held by the candidates constitute an opportunity for observation by the commissioned officers at the detachment. And an extensive Report of Observation on each candidate in his several jobs is prepared by the grading officer. Figure 1 is a reproduction of one of the forms used to grade candidates at the Special Training Detachment when they act as instructors. It is a convenient form which includes space for the name of the candidate rated, his OCS section and class, the subject taught and the date. The candidate is rated on 10 personal characteristics: Military Bearing and Appearance; Aggressiveness; Voice; Poise and Assurance; Application; Judgment; Initiative; Ability to Learn; Adaptability; and Personality. Each such category is broken down into several grades, 75 being considered the minimum satisfactory grade. For each category, several descriptive phrases are included and the observing officer need only check the particular phrase applicable to the candidate. Boxes for each category are also included to indicate the grade for each of the 10 personal characteristics. Space is provided for an average grade for the job done and signature of the observing officer. The reverse side of the Report of Observation, not shown in the reproduction, allows for a description by the commissioned officer of the capacity in which the candidate was acting and for any additional comments the officer might wish to include on the candidate's performance.

The various Reports of Observation on a given candidate are collated and at the completion of a job or series of jobs, the candidate is given a critique on his performance. This critique is, in a sense, a clinical interview between the commissioned officer and the candidate. In complete privacy, the candidate and officer discuss the former's work sample and the officer points out where the candidate has been successful or unsuc-

NAME	SECTION			CLASS			SUBJECT			DATE		
	50	55	60	65		70	75		80	85	90	95
MILITARY BEARING AND APPEARANCE	SLOVENLY CARELESS			COLORLESS NOT ERECT			NATURAL CLEAN CUT			INSPIRING POLISHED		
AGGRESSIVENESS	MEEK OVERBEARING			RESERVED IRRESOLUTE			ENERGETIC DETERMINED			FORCEFUL DYNAMIC		
VOICE	IRRITATING DEFECTIVE SPEECH			LIFELESS WEAK			PLEASANT EFFECTIVE			CONVINCING COMMANDING		
POISE AND ASSURANCE	EMOTIONAL IMMATURE			NERVOUS APOLOGETIC			STEADY DELIBERATE			MATURE DECISIVE		
APPLICATION	INDOLENT INDIFFERENT			INEFFICIENT PASSIVE			DILIGENT SINCERE			CONSCIENTIOUS ZEALOUS		
JUDGMENT	INDISCREET UNREASONABLE			INDECISIVE ILLOGICAL			TACTFUL DEPENDABLE			RELIABLE MATURE		
INITIATIVE	INERT SPIRITLESS			SLOW AVOIDS RESPONSIBILITY			RESOURCEFUL ENERGETIC			VERY RESOURCEFUL ZEALOUS		
ABILITY TO LEARN	DULL INAPT			SHORT-SIGHTED SLOW TO GRASP			RATIONAL DISCERNING			ALERT CLEVER		
ADAPTABILITY	INTOLERANT ARGUMENTATIVE			OPINIONATED RESISTS CHANGE			COOPERATIVE IMAGINATIVE			VERSATILE ADJUSTS RAPIDLY		
PERSONALITY	AFFECTED INSIPID			PHLEGMATIC UNRESPONSIVE			SINCERE ANIMATED			VIBRANT DYNAMIC		

AVERAGE GRADE

REPORTING OFFICER

FIG. 1. Sample copy of Report of Observation.

cessful, and attempts to reinforce those personal characteristics which are deemed satisfactory and to stamp out the undersirable characteristics. The critique serves as an informal situation in which the candidate may talk over his difficulties and receive guidance from the commissioned officer. It is at this time that the greatest amount of skill and resourcefulness are required on the part of the commissioned officer to shape the

development of the candidate according to the OCS conception of the good officer.

We have sketched the general organization and purposes of the Special Training Detachment. What have been the results, so far, of this type of training? Table 2 presents this information on classes 39–49.

From Table 2 we may see that 173 candidates were turned back for leadership training at the Special Training Detachment. Twenty-three of these or 14% of the entire group were failed at the detachment. Forty-one or 24% of the group were given the month's course in leader-

Table 2

Data on Commissions or Failures after Special Training Detachment Course

Class	Number Turned Back	Number Failed at STD		Number Failed After Field		No. Commissioned After Field	
		No.	%	No.	%	No.	%
39	20			7	35	13	65
40	28	10	36	4	14	14	50
41	23	5	22	3	13	15	65
42	8	1	12	4	50	3	38
43	15			1	7	14	93
44	14	2	14			12	86
45	20	2	10	6	30	12	60
46	20	1	5	8	40	11	55
47	12	2	16	5	42	5	42
48*	4			1	25	3	75
49**	9			2	22	7	78
Totals	173	23	13	41	24	109	63

* Class 48 Number turned back excludes 4 candidates who resigned from OCS either during STD training or immediately thereafter.

** Class 49 Number turned back excludes 1 candidate who resigned from OCS during STD training.

ship training but were not commissioned after participating in the field exercise. Typical comments recorded by the Faculty Board of officers for relieving these candidates were: "Relieved for leadership deficiency; lack of force and aggressiveness; timidity and lack of personal force; inability to exercise sound judgment; lack of capacity for teamwork; lack of adaptability." One hundred and nine or 63% of the entire group were commissioned after the field exercise. Thus, 63% of the group achieved success, as compared with a total of 37% who did not graduate. Now, if we consider simply the total number of men who were sent to the field exercise, we see that our group consisted of 150 candidates. Of this

number, 41 or 27% were unsuccessful candidates but that 109 or 73% of the group going through the field exercise were commissioned. Mention should also be made here of the three candidates who because of exceptional qualities were commissioned directly at the Special Training Detachment. These figures demonstrate even more dramatically the effectiveness of the detachment's training. Thus, we may conclude that the Special Training Detachment was a significant factor in the reclamation of those candidates who might otherwise have been washed out for lack of leadership qualities.

Summary and Conclusions

This paper describes the attempt of the Signal Corps Officer Candidate School at Fort Monmouth, N. J., to deal with an important classification of candidate failure,—lack of leadership qualities. The Special Training Detachment of this OCS has been discussed and data presented to demonstrate the success of the program in developing these qualities.

Received January 27, 1946.

A Note on Popular Pseudo-Psychological Beliefs in 1923 and in 1946 *

Howard P. Longstaff

University of Minnesota

Paterson [1] in 1923 made a study of the extent to which college students taking a course in applied psychology during the summer session at the University of Minnesota believed in certain pseudo-psychological systems of character analysis, along with a few prevalent false beliefs in psychology. He found that a surprisingly large percentage of these advanced students believed in these pseudo-scientific systems. Psychologists have done a considerable amount of work to combat these systems since 1923, especially in the beginning course in psychology, so it seemed pertinent to test the effectiveness of these efforts in reducting belief in these unscientific "ologies."

The present study was made during the second summer session of 1946 at the University of Minnesota. Advanced students taking two courses in applied psychology were questioned about the same systems previously investigated by Paterson. The first group of 53 subjects, taking the course Occupational and Vocational Psychology, were given the questionnaire used by Paterson. The second group of 124 students, taking the course Psychology in Personnel Work, were given a slight modification of this same questionnaire. The modification consisted of stating the nature of the system rather than merely giving its name as Paterson did, e.g., the first item in Paterson's questionnaire was "astrology" while in the modified form the first item was the statement, "The stars influence our behavior." This modification was used to test the assumption that students may or may not believe in the system to the same extent when only the name is used as when the essential tenet of the system is given. There was a slight tendency for the students to be more gullible on the modified questionnaire but the differences were so

* This article is a "prior publication," the author paying complete costs. The scheduled 80 pages per issue is thereby increased by the corresponding amount; thus the "early publication" of this article is a direct contribution to the subscribers of the *Journal of Applied Psychology* without handicap to those authors whose articles are accepted and printed in their regular turn.

[1] Paterson, D. G. A note on popular pseudo-psychological beliefs. *J. appl. Psyhol.*, 1923, 7, 101–102.

slight that chance could well account for them. Therefore, the data from the two groups have been combined for our consideration here.

Table 1 presents the comparative data. While the table is practically self-explanatory a few comments may be in order. We note that a significantly smaller number of students in 1946 believe in chiropractic, mental telepathy, physiognomy, or that gifted children are peculiar, while the differences between present beliefs and those in 1923 approach statistical significance in the case of spiritualism and phrenology. There

Table 1

Number and Per Cent of Students who Indicated Positive Belief in the Systems Listed in the Left-hand Column in 1923 and in 1946

	1923 Total N = 20		1946 Total N = 177			
	N	%	N	%	D	D/σD
Astrology	2	10	6	3	7	1.0
Chiropractic	10	50	15	9	41	4.0
Fortune telling	2	10	4	2	8	1.1
Graphology	5	25	18	10	15	1.5
Hypnotism	12	60	147	83	23	2.1
Memory systems	12	60	94	53	7	0.6
Mental telepathy	12	60	13	7	53	5.0
Spiritualism	5	25	2	1	24	2.4
Palmistry	1	5	3	2	3	0.6
Phrenology	5	25	0	0	25	2.5
Physiognomy	11	55	5	3	52	5.0
Psychotherapy	9	45	158	89	44	4.0
Absent treatment	1	5	2	1	4	0.8
Gifted children are peculiar	12	60	10	6	54	5.0
Psychology can determine what specific job a man is best fitted for	16	80	85	48	32	3.2

is also a decided drop in the per cent who believe "psychology can determine what specific job a man is best fitted for." This change, however, does not necessarily mean that present-day students do not believe in vocational guidance but rather a larger proportion of them have learned to reject the dogmatic statement that psychology can determine the *specific* job one should enter. Many students made this clear by commenting upon this question.

There is a smaller change in the number who believe in astrology, fortune telling, graphology, palmistry, and absent treatment. This smaller drop, however, is probably due to the fact that most of these systems were believed in by only a small proportion of the students in

1923. About 60 per cent believed in memory systems in 1923 and 53 per cent believe in them in 1946. By way of contrast, there is a significant increase in the percentage of present-day students who believe in hypnotism and psychotherapy. This finding probably reflects a greater emphasis on hypnotism and psychotherapy in the literature and in the content of the beginning course than was true in the early 1920's.

Considered as a whole, the data indicate that the beginning course in psychology is far more effective in combatting belief in pseudo-psychology than it was two decades ago. But there are still students who believe in purely quack systems. Since we find this among students in advanced psychology courses, there are probably a considerably larger number of people in general who still believe in these systems. So, psychology's task is not yet completed in dealing with this problem.

Received November 27, 1946.

Applied Research That Might Have Been *

Kenneth E. Clark

University of Minnesota

Psychologists working in and with the Navy Department during World War II worked on problems and projects covering an extremely wide range. That their contributions were exceedingly valuable to the Navy is indisputable. Equally indisputable, however, is the fact that many projects were of limited value or failed to result in improved utilization, training, or welfare of naval personnel.

Since many groups will use the war-time experiences of psychologists as a guide to peacetime work, some mediocrities of attainment may well be cited as a warning to all professional psychologists to be on their guard against some common errors that are frequently made. The following issues are illustrated by excerpts from recent research reports on Navy projects.

1. Is the problem one which can be solved as effectively by administrative changes or changed policies as by research? The following statement appeared in a recent report: "While the study was in progress the problem largely disappeared. Complements of Radarmen were increased for the types of ships serviced at Norfolk so that Radarmen strikers were no longer selected by the selection officers. As a result, no effort was made to develop a more satisfactory criterion to test unselected groups of men in the Destroyer Pool, or to try other measures, and no formal report was ever made." It appears to the writer that this problem should never have been attacked as a research problem, since another solution was so readily available.

2. Is the problem of sufficient significance to warrant its study? The following candid statement is quoted: "The Project's report on 'A Procedure for Sorting the Enlisted Personnel Qualifications Card' has received tardy attention. According to the last information received by the Project, the proposals in this report were being studied by Service Forces Pacific with a view to determining whether the proposals could be coordinated with certain modifications of fleet personnel accounting procedures then under consideration. There has been no evidence of practical effectiveness of this report." This excerpt indicates a situation which recurred frequently throughout the war when psychologists at-

* The opinions or assertions contained herein are the private ones of the writer and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large.

tacked the problem first, and its significance second. Here is another example of this sort: "The tests developed by the Project for use in the Amphibious Training Base at Coronado, California, were discarded shortly after the Project representative left the base. The reason given for this was a reduction in staff and the consequent lack of personnel to execute the testing program. Whatever the cause, the fact that the tests were not continued in use means that the Project's effort bore only temporary fruit."

3. Are problems accepted for study on the basis of their significance for improved practices or merely because they are suggested for study? Certain problems which confronted the armed services were never attacked by those personnel best qualified to study them because of insufficient aggressiveness on the part of professional personnel. The following excerpt illustrates this point: "The Project has been concerned about certain aspects of the Navy's quality system for the assignment of recruits to service schools but due to the lack of formal authorization has refrained from undertaking rigorous research into the matter or expressing its viewpoint in a formal communication." Yet, while such projects were not undertaken, other projects of limited or temporary value were accepted for study.

4. Are recommendations stated in sufficiently broad terms to cover all factors which may affect results, or are they limited in value by a narrow construction of the field of work for psychologists? Possibly because of high standards of academic integrity, professional psychologists have usually avoided attacking those aspects of problems lying too far outside their own areas of specialization, preferring to plan and accomplish cooperative research projects with other professional groups. This reluctance to exceed the bounds of one's own field is probably being carried too far when such observations as the following are not even referred to in recommending changes on the basis of research studies.

Observation: *"Availability of the Enlisted Personnel Qualifications Card Aboard Ship:* A Project member examined the files of Enlisted Personnel Qualifications Cards aboard three destroyers. In spite of the fact that the crews of these vessels had been formed only two to four months earlier at a Precommissioning Center, two related types of irregularities were discovered: (a) Qualifications Cards were missing for 34 men in a combined crew of 1,000; (b) 23 Qualifications Cards were found for men who were not on board ship. . . . Another deficiency of the files of Qualifications Cards was the frequent absence of Basic Battery test scores for men whose Cards were prepared at the Precommissioning Center. (Qualifications Cards were available for only about 20 per cent of such men.) It was further observed that very few of the officers on

board were aware that the information on the Q-Card could be used in assigning men to billets."

Recommendations: "*Improvement of Personnel Records:* Efficient personnel procedures are impossible without adequate records. Efficient records are impossible without standardization of the terminology used to express qualifications, experience, preferences, etc. Such standard terminology exists only in partial degree in the Navy today. Another source of difficulty in personnel records is the mere lack of space to record the increasing amount of information which an improved personnel system supplies. Thus, there is inadequate room on the Enlisted Personnel Qualifications Card (NavPers 609) for all of the 'special tests' which a recruit may have. In general, too, there is no provision for a possible future re-evaluation of the recruit on the basis of his experience, or of possible re-tests. . . . Finally, it is desirable that, wherever possible, personnel records be made more conveniently usable; Report No. 15 ('A Procedure for Sorting the Enlisted Personnel Qualifications Cards.' Project Report No. 15 OSRD Report No. 4689, January 22, 1945.) of this Project provides an example of what might be done in this direction."

5. Do reports prepared for the use of management maintain standards for reporting research results which are required in normal professional publication? About June of 1944, the Navy adopted for use in classifying enlisted personnel the Bausch and Lomb Orthorater. This instrument is of use in measuring acuities, phorias, and stereopsis, and has demonstrated in industry its value for discriminating between high and low production workers. Several different Navy research groups were interested in the Orthorater, and differed widely among themselves regarding its value for Navy use. In March, 1945, a report on the selection of Rangefinder Operators was published, recommending use of the Orthorater with either of the following sets of requirements:

"After investigation of many possible combinations of the tests at various cut-offs, it appears that in terms of selections and instruction cost and other practical considerations, the following two batteries yield the most satisfactory results:

1. Battery No. 1.
   a. General Classification Test: 50 and above;
   b. Arithmetic Test: 45 and above;
   c. Orthorater Visual Acuity worse eye score: 9 or greater;
   d. Orthorater Vertical Phoria score: not less than 4 nor more than 7 (Hyperphoria not greater than ½ prism diopter);
   e. Orthorater Lateral Phoria score: not less than 4 nor more than 12 (Esophoria not greater than 3½ prism diopters, Exophoria not greater than 4½ prism diopters);
   f. Multiple Projection Eikenometer Test: 58 or greater.

2. Battery No. 2
   a. General Classification Test: 45 and above;
   b. Arithmetic Test: 55 and above;
   c. Orthorater Visual Acuity worse eye score: 9 or better;
   d. Orthorater Vertical Phoria score: not less than 4 nor more than 7 (Hyperphoria not greater than ½ prism diopter);
   e. Orthorater Lateral Phoria score: not less than 4 nor more than 12 (Esophoria not greater than 3½ prism diopters, Exophoria not greater than 4½ prism diopters);
   f. Orthorater Depth score: 5 or greater."

"C. The results of both final batteries have been tested for statistical significance and more than satisfy the 5 per cent level of confidence. This is true for Class 6 (N = 56), for Class 7 (N = 90) and for the two classes combined (N = 146). Consequently, similar results can be expected with reasonable certainty for similar Navy populations."

This report and its recommendations, based on analysis of the records of 146 students, presented neither the original data nor the statistics computed from these data, except in terms of specific cutting scores. That these scores were assigned to maximize the differentiation between good and poor students by using chance differences is suggested by the changes in cutting scores used for the General Classification Test and the Arithmetic Test from Battery No. 1 to Battery No. 2. At any rate, insufficient data were presented to permit comparisons with results of other studies, or to permit further analysis.

6. Can this study be undertaken and completed in time to make its results of value? A statistical evaluation of the Navy's basic classification battery, Form I, presented distributions of test scores, data on reliability, on inter-correlations of tests, and made recommendations regarding the types of items and the kinds of tests which should comprise such a test battery. Form I of this test battery was installed for Navy-wide use on June 15, 1943. The first revision, Form II, was installed January 1, 1944, followed by Form III, which was installed January 10, 1945. The final report of the evaluation of the Form I battery was published March 10, 1945. In the actual development of the revisions to Form I, much of the information published in the report was available for use; yet to most Naval personnel, the lateness of the report's publication produced feelings of exasperation at the apparent slowness of research.

7. Are all aspects of the problem, and its relationship to operations and general practices fully understood by the investigator? A report dated September 7, 1945, presented data indicating the relative merits of interviewers' recommendations and test scores for predicting success in Navy service schools. Interviewers' recommendations were made using

a quality code reflecting the man's aptitude and adaptibility for schooling and a school code indicating the kind of school in which he was judged most likely to succeed. The findings in favor of the test scores for predicting success become meaningless when one discovers that the sample used for this study was processed at naval training centers prior to January 1, 1944, yet the procedures for using school and quality codes in the manner being evaluated were not installed until March 1, 1944. This fact, plus misstatements in the report which indicated lack of knowledge of the methods used in filling quotas, made the entire report seem ridiculous to personnel familiar with recruit assignment procedures.

It should not be inferred from the emphasis in this article that wartime psychological research was generally ineffective or remiss. Wartime research was necessarily hurried; every problem carried a high priority either because of our national military situation, or because of the pressure of large numbers of personnel to be processed. That their results were generally excellent, and that the applications of research findings materially aided the war effort cannot be over-emphasized.

Likewise, it should not be inferred from the preceding examples that the worst examples of wartime research deficiencies have been cited. Rather, those reports which were sufficiently candid and forthright to indicate their own limitations have been made to suffer out of proportion to their sins. Omitted is any mention of studies where errors or deficiences are not obvious, and those which neglected to mention the degree to which their recommendations were used or ignored.

Received March 21, 1946.

Book Reviews

Benge, E. J. *How to make a morale survey, a manual of procedures.* Deep River, Conn.: National Foremen's Institute, 1941. Pp. 64

The title of of this book implies that it is sufficiently detailed to enable the reader to make a morale study. This purpose is adequately fulfilled in only some respects. The discussion of the nature of morale is handled very satisfactorily in a non-technical manner. The author has a knack for drawing analogies which will appeal to executives in business and industry and which will convince such persons of the value of morale surveys. He is to be commended for emphasizing throughout the book that a morale survey should not be made unless management is willing to face the facts and act in accordance with the findings. Many practical hints, primarily on a non-technical or semi-technical level, are given in sections on constructing morale questionnaires and conducting the survey. Various methods for determining morale are enumerated, but are given little or no discussion except for the multiple choice questionnaire method (which is recommended) and the use of scaled statements.

The section on analyzing results is weak. Inconsistently, the author warns against comparing a morale index on one item with that of another (because arbitrarily assigned values may not be equated), but in numerous instances he has combined results from various items in order to secure an overall index. The only method mentioned for determining values of scaled statements is based on ranks which are handled as linear measures. The weakness of summing scale values of statements endorsed by employees is brought out, but no mention is made of using a median or mean scale value or algebraically summing positive and negative scale values as a possible remedy. The author admits theoretical statistical objections to this method of scaling statements, but states the method works well in practice. The reviewer is more inclined to agree with a later statement that when a company has gone to the trouble of making a morale survey the statistical analysis should not be slighted.

The book was not intended for nor apt to be valuable to psychologists. Its greatest use will be to impart general procedural information to executives. It is unfortunate that the author has included sections such as use of IBM equipment in handling data, for these are apt to give the impression that the book was designed as a manual which safely permits morale surveys to be made by untrained persons. The author does recommend, however, that surveys be conducted by consultants,—

who presumably would have a better background than could be obtained from this book.

Some psychologists may have difficulty in justifying to executives their use of more rigorous techniques than those advocated here. Unfortunately, errors of over-simplification (e.g. treating ranks as linear measures) may result in executives believing that persons who advocate more complicated techniques are "theoretical" or "impractical."

The page size is eleven by eight and one half inches, and pages are in a loose leaf, three ring, leather binder. In addition to the sixty-four numbered pages there are twenty-two pages of figures.

Clifford E. Jurgensen

Minneapolis Gas Light Co.

Brighouse, Gilbert. *The physically handicapped worker in industry.* Pasadena: California Institute of Technology, 1946. Pp. 54. $2.00.

The war years offered the physically handicapped the greatest opportunity in history to demonstrate that, if properly selected and placed, they can perform as efficiently as physically normal individuals. This investigation includes an objective comparison of the characteristics and performance of 1,815 physically handicapped persons with 538 non-disabled persons employed at the Lockheed Aircraft Corporation during 1944–45.

The disabled group—1,117 men and 698 women—included twenty-seven types of disablement which were readily classified into eight major groups. The physically normal or control group consisted of 338 males and 200 females selected at random from the records of all persons currently employed.

The comparison of personal characteristics of the two groups included age, work shifts, employment in hazardous areas, martial status, home ownership, educational achievement, previous occupations and character investigation. As a group, the physically handicapped were no different from the control group and possess characteristics similar to those of non-disabled persons. In spite of the identification of disablement by the Medical Department, the same percentage of disabled as non-disabled were placed in hazardous work areas.

The employment efficiency of an individual or a group cannot be measured in terms of a single criteria. Therefore, to determine the effectiveness of the physically handicapped, comparisons were made of a number of variables, the majority of which were relatively susceptible to objective treatment. These included varieties of work, levels of responsibility, rate of promotion, frequency of job change, length of service, approvals for rehire, supervisory ratings, absences, participation in group

insurance program, violation of company rules, special commendation, grievances and the results of a morale questionnaire.

Such differences as did occur were not of statistical significance. Those in favor of the handicapped group included less labor turnover, greater stability on job, less readiness to quit, more rapid advancement through lower levels jobs, less unexcused absenteeism, greater willingness to remain on "dead end" jobs, and slightly higher morale. Obvious handicaps tended to be an incentive to higher morale on the part of the physically normal.

In the negative direction, the handicapped were less likely to be promoted into a position of responsibility, slightly more accident prone, more frequent absences because of illness, less likely to receive superior ratings from supervisors and special commendations. The slightly higher accident rate may no doubt have been due to placement of 23 per cent of the disabled in hazardous areas. It was the opinion of the safety engineers that the handicapped were less likely than the non-disabled to have accidents because of their own carelessness or inexperience.

This study adds a variety of objective evidence to the principle that the characteristics of the physically handicapped are similar to those of the non-disabled and that if properly selected and placed in employment can perform as efficiently as any other group of workers. In quality, the study is decidedly superior to the majority which have appeared during the war period. It should be of particular interest to employers, persons engaged in personnel programs, and those engaged in the vocational rehabilitation of disabled civilians and veterans.

D. H. Dabelstein

Office of Vocational Rehabilitation,
Washington, D. C.

Faist, Kenneth L., and Newkirk, Stanton M. *Job safety training manual.* Deep River, Conn.: National Foremen's Institute, Inc., 1944. Pp. 52. $5.00.

This is a manual designed for leaders of safety training programs. The technique is basically the same as that of the "J" programs of T. W. I. This is the JST to be added to JIT, JMT and JRT.

The program is divided into five two-hour sessions. The procedure that is taught is to (1) analyze the job, (2) spot the danger, (3) act to correct, and (4) check back. The primary emphasis is upon individualism in training,—foremen are taught to analyze individual men and jobs into job steps and KNO points (or activities and places that are hazardous), and to take steps to prevent those accidents.

This is all brought out in the first session where the right and wrong methods of showing a new man his job are demonstrated. Later sessions

apply the four steps to older employees on new jobs and to older employees who are doing their customary work in an unsafe manner. The manual contains detailed material for the leader to use in the five sessions.

Some suggestions for post-session continuation of the training program are described at the back of the manual. Sample booklets and forms for supervisors and employees are included. Three appendices contain: (1) a form letter for obtaining the reactions of participants to the program, (2) a description of the safety committee work at the Perfect Circle Company and at one General Electric plant; and (3) a brief annotated bibliography of references and sources.

This program follows very closely the methods of the "J" programs and should be equally as successful. The procedure here differs from most safety training programs in the commendable emphasis on individual workers and work-places. Safety education is too often a matter of bulletin boards, posters, and unanalyzed figures. This program is also shorter, more intensive, and probably more interesting than most.

The one point that might be criticized severely is the omission of psychological aspects of safety. Although there is adequate treatment of specific habits and motor skills, there is no consideration of the more general aspects of behavior. "Accident-proneness",—or, as it would better be called, "accident-repetition",—is not mentioned. Foremen are not told to analyze the personalities, emotional habits and swings, personal problems, etc., of their subordinates. There is no suggestion that foremen collect accident statistics, or near-accident statistics, analyze these, and attempt to relate them to other personnel data.

In general this manual should make for better safety training. Further improvements should result from adding some psychological and statistical control procedures to the program, particularly from a discussion of "accident-repeaters." Perhaps this could be achieved most conveniently by devoting one meeting exclusively to that and dropping one of the demonstrations of analysis into job steps and KNO points.

Harold F. Rothe

Stevenson, Jordan and Harrison, Inc.,
Chicago, Ill.

Shartle, Carroll L. *Occupational information: Its development and application.* New York: Prentice-Hall, Inc., 1946. Pp. xiv + 339. $3.50.

"This book was written to meet the apparent need for an introduction to the development of occupational information and a description of the uses of such information in problems in industry, government, education, and community agencies. It is felt that the individual who develops or uses such information should have a background of certain techniques

and should be acquainted with their uses not only in his particular specialty but also in related fields. Knowing the limitations as well as the advantages of occupational information is likewise useful."

With this introductory statement of purpose, Dr. Shartle, who is uniquely qualified by training and experience to author this book, proceeds to rescue the field of occupational information from the hands of well intentioned but inadequately equipped teachers of ninth grade Occupations courses and their cohorts and to give it the status befitting its acknowledged importance in a variety of personnel situations. As a presentation of an organized body of knowledge and a statement of methodology the appearance of this volume heralds the coming of age of what Kitson has tentatively labelled for want of a better term the "Science of Occupationology." With the publication of this book it is no longer possible for colleges and universities to ignore demands for instruction and training in this area. Even academic psychologists will find it to their interest to examine this book carefully if for no other reason than the fact that their students are, in increasingly great numbers, finding employment in fields which make familiarity with this type of material a basic requirement. As a text, *Occupational Information* will find its greatest usefulness in the training of vocational counselors and business and industrial personnel workers. Instruction may be lodged administratively in any of several units depending upon local circumstances. In at least two instances, the author's and the reviewer's courses, all of the materials included here are a part of the personnel psychology curriculum.

The chapter headings are descriptive of the content. Following the introductory chapter in which occupational information is defined to include accurate and usable information "about jobs and occupations . . . about industries and processes insofar as such information is directly related to jobs . . . about occupational trends and the supply and demand of labor," the remaining ten chapter headings are Obtaining Occupational Information, Describing Jobs and Occupations, Classifying Jobs and Occupations, Using the Dictionary of Occupational Titles, Occupational Families, Entry Fields of Work, Military to Civilian Occupations, Industries and Patterns of Occupations, Jobs for the Handicapped, and Current and Future Occupational Opportunities. The reviewer found the chapter Jobs for the Handicapped particularly rewarding. Military to Civilian Occupations is probably the least important both because of the time factor and because of the limitations on the use of this material which the author has wisely pointed out.

The following characteristics may be counted among the outstanding virtues of the book: (1) definition and standardization of terminology, (2) coverage of a representative range of materials, techniques, and pro-

cedures, (3) use of a wealth of illustrative materials, (4) suggestions for practical applications and adaptations particularly to local situations, (5) inclusion of helpful exercises including a set of training materials used in the U. S. Employment Service, and (6) listings of source materials.

Because of the importance of this publication in defining the field of occupational information it is necessary to examine critically its basic orientation. The reviewer admits to a difference of opinion regarding the advisability of disassociating worker and job analysis such as is done in this volume. This separation is indicated by the following quotations. "Occupational information does not include the study of abilities, aptitudes, and other characteristics of individual workers" (pp. 1–2) and "One must be careful to differentiate job analysis from worker analysis" (p. 13). If one accepts, as is generally done, the viewpoint that counseling and placement activities are facilitated when job and worker analysis are made in the same terms the present separation is regrettable. Although brief mention is make of intelligence and interest classifications of occupations and the possibility of forming job families on the basis of measured worker characteristics this whole approach is rather summarily dismissed from further consideration. The major emphasis throughout is upon a "duties and working conditions" type of occupational information. Worker characteristics are a part of this type of occupational information only in the form of arbitrary employer requirements such as age, education and special training or as estimates by the job analyst of the amounts of various characteristics required for successful performance of the job such as keenness of vision or memory for details. Both of these types of information are subject to serious drawbacks. The first perpetuates all the errors of arbitrary and often unrealistic requirements made by employers. The usefulness of the second type is limited by the fact that job analysts may be in serious error in estimating the importance of a given characteristic. For example, clerical aptitude may be the important differentiator between successful and unsuccessful order fillers in a warehouse rather than the more "obvious" manual dexterity estimated by a job analyst to be required in large amount. The reviewer recognizes the practical difficulties in the way of obtaining information on measured worker characteristics. However, a book which will have such an important part in influencing thinking in this field should treat not only of present practices but should also point out future lines of development. The integration of job and worker analysis whereby both sets of facts have a certain common core provided by objectively determined worker characteristics is one such line of development.

Two additional points might be noted. The person unfamiliar with the materials of this field may be mislead unintentionally into thinking

that a wealth of appropriate and sound occupational information is available. This book is really a catalogue of existing "best practices." Actually many of the materials illustrated do not exist in sufficient variety to be very useful. For example, the job and industrial families published by the Employment Service are limited primarily to the important war occupations and industries. The same is true of their excellent series of Individual Occupational Descriptions. Publications of the nongovernmental agencies often have a wider coverage but, in some instances, do not meet reasonable standards of quality and reliability. One of the important values of Dr. Shartle's book is that it provides an excellent background for evaluating such materials. Again, the person new to this field may wonder at the prominence of governmental materials in this treatment. This reflects not only the author's long association in a top administrative capacity with the United States Employment Service but also the fact that the kind of research and developmental activity responsible for the production of really comprehensive occupational information requires large scale operations which are possible only through the cooperation of national agencies with local units. One is tempted to draw the moral that consumers of such materials will do well to support measures designed to increase the effectiveness of such agencies or actually to prevent their emasculation.

Occupational Information is authoritative and timely. Dr. Shartle and the publishers are to be congratulated upon the production of a comprehensive survey of an increasingly important personnel tool.

Arthur H. Brayfield

Long Beach City College, California

Erickson, C. E. and Happ, M. C. *Guidance practices at work.* New York: McGraw-Hill, 1946. Pp. 325.

This is one of the first of McGraw-Hill's new "Practical Guidance Series" edited by Franklin A. Zeran. It is a compilation of descriptions of guidance activities, procedures, and techniques gathered by the authors from a large number of school officials and teachers in a great many different schools, with brief introductory and summary remarks by Erickson and Happ. As the authors aptly put it in opening their final chapter (p. 280), "Many descriptions of specific practices have been described."

The book is presumably designed to show practitioners and students of guidance how the various aspects of a guidance program are carried on in a variety of schools, and to serve as a source of ideas for improving local practice. Unfortunately, the quality of the material included in this compilation is very uneven, much of it being descriptive on a superficial or back-slapping level (e.g. page 304: "The outstanding accomplishment,

for which much credit is due the placement adviser . . . "). If this material actually represented the *best* practices in guidance, it would be a cause of real discougarement for those who have been training counselors and to those who, out in the field, have striven to develop sound programs. The frequently meaningless comments of the general authors do little to improve the impression caused by the poor selection of material.

One looks in vain for descriptions of such well-developed programs as those in Pasadena, Providence, Baltimore, Rockland County (N. Y.) and Cincinnati, for such names as those of Mary Corre, Leona Buchwald, and Margaret Bennett. An emphasis on new programs and new names would be helpful if they had something to contribute. Instead, some of the practices are actually unsound, for example that on page 200 entitled "In Chicago High Schools Every Teacher a Counselor." If, as stated there, "every teacher . . . must now learn to become (a) personality guide as well" by devoting an hour each day to individual counseling, one trembles for the pupils. Any group of teachers employed as subject-matter specialists is sure to include a substantial proportion who are not personally qualified for, interested, or trained in counseling. Some of the practices described are original and worthy of becoming more widely known, for example the "Preview to Nursing" exploratory program described in thirteen lines on page 173. But two or three journal articles could more effectively have brought these to the attention of the profession.

Those of us who teach courses in guidance would perhaps do well to ascertain the reaction of our students to a book such as this. If the students find it helpful, we may be made aware of a great gap between the level at which we expect our students to function and the level at which they do actually operate in the schools. In that case, a need for more emphasis on the "how-to-do" type of instruction is indicated. It is to be hoped that it will be more selective and will lead the students to higher levels of study and practice than the material in this book.

Donald E. Super

Department of Guidance,
Teachers College, Columbia University

New Books, Monographs, and Pamphlets

Books, monographs, and pamphlets for listing and possible review should be sent to Donald G. Paterson, Editor, Department of Psychology, University of Minnesota, Minneapolis 14, Minnesota

How to conduct consumer and opinion research. Albert B. Blankenship. New York: Harper & Brothers, 1946. Pp. 314. $4.00.

Out of uniform. Benjamin C. Bowker. New York: W. W. Norton & Co., 1946. Pp. 259. $2.75.

Educational sociology. F. J. Brown. New York: Prentice-Hall, Inc., 1946. Pp. 736. $3.75.

Employment, rehabilitation, and veteran adjustment. Paul S. Burnham. Chicago: Public Administration Service, 1946. Pp. 27. $1.00.

How to supervise people in industry. Elliot Chapple and Edmond Wright. Deep River, Conn.: National Foremen's Institute, Inc., 1946. Pp. 123. $2.50.

Psychology in action. Joseph Clawson. New York: The Macmillan Co., 1946. Pp. 289. $4.00.

Executive ability. Glen U. Cleeton. Yellow Springs: The Antioch Press, 1946. Pp. 540. $4.50.

The application of the Rorschach test for young children. No. 23. Child Welfare Monograph Series. Mary Ford. Minneapolis: University of Minnesota Press, 1946. Pp. 114. $2.00.

Dream mates. Leonhard Frank. New York: Philosophical Library, 1946. Pp. 245. $3.00.

Foundations of the measurement of values. Bertha B. Friedman. New York: Bureau of Publications, Columbia University, 1946. Pp. 227. $2.75.

The psychology of adolescence. Karl C. Garrison. New York: Prentice-Hall, Inc., 1946. Pp. 355. $4.65.

Psychology in human affairs. J. Stanley Gray. New York: McGraw-Hill Book Co., Inc., 1946. Pp. 646. $3.75.

Creative old age. Clare de Gruchy. San Francisco: Old Age Counselling Center, 1946. Pp. 143. $2.75.

The nature and conditions of learning. Howard L. Kingsley. New York: Prentice-Hall, Inc., 1946. Pp. 579. $4.50.

Managing your mind. S. H. Kraines and E. S. Thetford. New York: The Macmillan Co., 1946. Pp. 374. $2.75.

Textbook of abnormal psychology. Carney Landis and M. Marjorie Bolles. New York: The Macmillan Co., 1946. Pp. 576. $4.50.

Intellectual status at maturity as a criterion for selecting items in preschool tests. No. 21. Child Welfare Monograph Series. Katharine M. Maurer. Minneapolis: University of Minnesota Press, 1946. Pp. 166. $2.50.

Human factors in air transport design. Ross A. McFarland. New York: McGraw-Hill Book Co., Inc., 1946. Pp. 670. $6.00.

The chemical senses. R. W. Moncrieff. New York: John Wiley and Sons, Inc., 1946. Pp. 424. $4.50.

Agnosia, apraxia, aphasia: their value in cerebral localization. J. M. Nielsen. New York: Paul B. Hoeber, 1946. Pp. 292. $5.00.

Youth in trouble. Austin L. Porterfield. Fort Worth: The Leo Potishman Foundation, 1946. Pp. 135. $1.50.

The relation of parental authority to children's behavior and attitudes. No. 22. Child Welfare Monograph Series. Marian J. Radke. Minneapolis: University of Minnesota Press, 1946. Pp. 123. $2.00.

Modern clinical psychology. T. W. Richards. New York: McGraw-Hill Book Co., Inc., 1946. Pp. 331. $3.50.

The nature of deterioration in schizophrenic conditions. David Shakow. New York: Coolidge Foundation, 1946. Pp. 88. $3.00

The human frontier. Roger J. Williams. New York: Harcourt, Brace and Company, 1946. Pp. 314. $3.00.

Interviewing procedure. Catalog No. L 24.2: In 88. Washington, D. C.: Superintendent of Documents, Government Printing Office, 1946. Pp. 25. $.10.

Efficiency of impaired worker. Federal Security Agency. Washington, D. C.: Office of Vocational Rehabilitation, 1946. Pp. 12.

The Carnegie Foundation for the Advancement of Teaching. Forty-First Annual report. New York: The Carnegie Foundation for the Advancement of Teaching, 1946. Pp. 160. Gratis.

Scientific Reports on Aviation Psychology, Washington, D. C.: Superintendent of Documents, Government Printing Office, Pre-publication deposit $35.00 for all reports:

The aviation psychology program in the Army Air Forces. Report No. 1. John C. Flanagan.

The classification program. Report No. 2. Philip H. DuBois.

Research problems and techniques. Report No. 3. Robert L. Thorndike.

Apparatus tests. Report No. 4. Authur W. Melton.

Printed classification tests, parts I and II. Report No. 5. J. P. Guilford.

The AAF qualifying examination. Report No. 6. Frederick B. Davis.

Motion picture testing and research. Report No. 7. James J. Gibson.
Psychological research on pilot training. Report No. 8. Neal E. Miller.
Psychological research on bomardier training. Report No. 9. Edward H. Kemp.
Psychological research on navigator training. Report No. 10. Launor F. Carter.
Psychological research on flexible gunnery training. Report No. 11. R. N. Hobbs.
Psychological research on radar observer training. Report No. 12. Stuart W. Cook.
Psychological research on flight engineer training. Report No. 13. John T. Dailey.
Psychological research on problems of redistribution. Report No. 14. Frederic Wickert.
The psychological program in AAF convalescent hospitals. Report No. 15. Sidney W. Bijou.
Psychological research on operational training in the continental air forces. Report No. 16. Meredith P. Crawford.
Psychological research in the theaters of war. Report No. 17. William M. Lepley.
Records, analysis, and test procedures. Report No. 18. Walter L. Deemer.
Psychological research on equipment design. Report No. 19. Paul M. Fitts.

Erratum

Through an over-sight, the Table of Contents for Volume 30, published in the December 1946 issue, failed to include the following item:

Harris, D. H. **Questionnaire and Interview in Neuropsychiatric Screening** 644